Technical Report 1042

Tacit Knowledge in Military Leadership: Supporting Instrument Development

Joseph A. Horvath and Robert J. Sternberg
Yale University

George B. Forsythe, Patrick J. Sweeney, and Richard C. Bullis
United States Military Academy

Wendy M. Williams and Martin Dennis
Yale University

May 1996

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Yale University

Technical review by

Michael Drillings
Clinton B. Walker

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-STP, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (from. . . to)
1996, May	Interim	October 1993-November 1995

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Tacit Knowledge in Military Leadership: Supporting Instrument Development	MDA903-92-K-0125
	5b. PROGRAM ELEMENT NUMBER 0602785A
6. AUTHOR(S) Joseph A. Horvath and Robert J. Sternberg (Yale University), George B. Forsythe, Patrick J. Sweeney, and Richard C. Bullis (USMA), and Wendy M. Williams and Martin Dennis (Yale University)	5c. PROJECT NUMBER A790
	5d. TASK NUMBER 1111
	5e. WORK UNIT NUMBER C03
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Yale University Department of Psychology P.O. Box 11A, Yale Station New Haven, CT 06520-7447	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-RM 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Technical Report 1042

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: Trueman R. Tremble, Jr.

14. ABSTRACT *(Maximum 200 words)*:

Army officers, drawn from three different institutional settings, evaluated tacit-knowledge items to support development of tests of tacit knowledge for military leadership. Items were identified that differentiated leaders varying in leadership experience (at a level of leadership) or in rated leadership effectiveness. Subject-matter experts also sorted the items into categories of like items. Analyses of the categories identified dimensions that described the organization of the items. Item discrimination and organizational location will be used to select materials for construction of tests of tacit knowledge for each three levels of Army Leadership: platoon leader, company commander, and battalion commander.

15. SUBJECT TERMS

Tacit knowledge Leader knowledge Leadership Leader effectiveness

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	90	

Technical Report 1042

Tacit Knowledge in Military Leadership: Supporting Instrument Development

Joseph A. Horvath and Robert J. Sternberg
Yale University

George B. Forsythe, Patrick J. Sweeney, and Richard C. Bullis
United States Military Academy

Wendy M. Williams and Martin Dennis
Yale University

Leader Development Research Unit
Trueman R. Tremble, Jr., Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

May 1996

Army Project Number
2O262785A790

Personnel Systems and
Performance Technology

FOREWORD

A primary mission of the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is to enhance military readiness through programmatic research that supports the effective performance of Army leaders. To accomplish this, ARI and the United States Military Academy (USMA) established the Center for Leadership and Organizations Research (CLOR) at USMA to conduct research as part of ARI's research program in the areas of organizational leadership and leader development, education, and training. The research here is part of the ARI exploratory development research program formulated and undertaken by the CLOR.

This report is the third product of a project jointly undertaken by researchers at USMA and at Yale University. The overall objective of the project is to test the applicability of a theory of tacit knowledge to military leadership. Previous research has shown that tacit knowledge, acquired through practical on-the-job experiences, is related to executive and managerial effectiveness in civilian organizations.

This report examines specific items of tacit leadership knowledge that commissioned Army officers acquired through practical experiences. This examination identified knowledge items that differentiated leaders who varied either in leadership experience or in rated leadership effectiveness. These differentiating items will be used to construct tests of the tacit leadership knowledge of platoon leaders, company commanders, and battalion commanders.

If the tests prove to be valid, this research will have practical implications for leader development. In particular, the tests will provide means for measuring the tacit knowledge of military leadership. The validation research will also indicate the likely worth of greater emphasis, in the Army's leader development system, on operational assignments and experiences for knowledge acquisition.

ZITA M. SIMUTIS
Deputy Director
(Science and Technology)

EDGAR M. JOHNSON
Director

TACIT KNOWLEDGE IN MILITARY LEADERSHIP: SUPPORTING INSTRUMENT DEVELOPMENT

EXECUTIVE SUMMARY

Research Requirement:

To support the development of assessment instruments based on (a) the theory and methods of tacit-knowledge research, and (b) substantive leadership knowledge acquired from Army officers during an earlier phase of the project.

Procedure:

Tacit-knowledge items were presented to Army officers in three institutional settings. These officers provided rating and sort data, which was analyzed in a variety of ways. Analyses were directed toward (a) identification of items with the greatest promise for incorporation into externally valid tacit-knowledge tests, and (b) derivation of a knowledge-space representation to be used in developing generalizable, structurally valid tacit-knowledge tests.

Findings:

An explanatory model of tacit knowledge was proposed and elaborated. This model constitutes a high-level task model for interpreting scores on tacit-knowledge items and for supporting the validity of score interpretations. Promising knowledge items were identified for use in future instrument development--items for which ratings of item quality were associated with exogenous criterion variables. The structure of the tacit-knowledge space was derived, based on subjects' sorting of knowledge in the sample, for use in future instrument development.

Utilization of Findings:

Findings from the current study were described in terms of a unified validity framework. The unified framework was intended to (a) show how the current findings will support the long-term goal of valid, useful tacit-knowledge tests, and (b) set the stage for future content selection, question construction, test construction, and test validation.

TACIT KNOWLEDGE IN MILITARY LEADERSHIP: SUPPORTING INSTRUMENT DEVELOPMENT

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

TACIT KNOWLEDGE IN MILITARY LEADERSHIP: SUPPORTING INSTRUMENT DEVELOPMENT

Introduction

A battalion commander whom we interviewed during the course of our research tried to explain to us how he goes about developing his subordinate (company) commanders. One method that he has found useful is to speak with the soldiers in his battalion about their own job descriptions, about the hours they work, about how they perceive the training they engage in, and about how information gets disseminated to them through the chain-of-command. The battalion commander we interviewed has found that, by asking these particular questions of soldiers in his battalion, he can identify areas of weakness in the command of particular companies and can use this information to structure his efforts to develop the captains who lead those companies. Consider several additional facts about this example of leadership knowledge.

First, note that the battalion commander's knowledge about how to identify the developmental needs of subordinate leaders was, according to his own report, acquired through experience as a commander rather than through formal teaching. For example, the particular set of questions that this commander asks soldiers in his battalion is nowhere set out in Army leadership doctrine. Nor is the way in which he phrases his questions or the way in which he approaches soldiers for this purpose. In short, the commander's knowledge about how to identify the developmental needs of his subordinate leaders (of which we have here described only a fraction) was based upon his own experience as a commander.

Second, note that the knowledge in question was not readily articulated by the battalion commander we interviewed. That is, he did not sit down and, unprompted, tell us "in order to develop subordinate leaders you must speak to their soldiers directly..." Rather, the knowledge described above came out of a "story telling" exercise in which the battalion commander was encouraged to tell the interviewers a story about one of his own leadership experiences and, with prompting from the interviewers, reflected upon the exact nature of the lessons learned from that experience. In short, the battalion commander's knowledge about how to develop subordinate leaders was knowledge that he may not have known he possessed--at least until he was encouraged to describe and reflect on his experience.

Finally, consider that the knowledge described above is of a type that may be expected to have consequences for leadership effectiveness--that of the subordinate leaders being developed and, by extension, of the battalion commander who leads them. Thus, if the knowledge described above is "good"--if it promotes the development of more effective company commanders--then we may say that the battalion commander is a better leader for having developed this method. Conversely, if the knowledge is "bad"--if it causes company commanders to be undermined in front of their soldiers--then we may say that the battalion commander is a poorer leader for having derived this lesson from his experience.

The importance of practical, experience-based knowledge of leadership is well recognized. United States Army training doctrine specifies that operational assignments, jobs within the Army, function not only to accomplish specific missions but also to develop the leadership capacity of job incumbents. Development through operational assignments is one of the three "pillars" of Army leadership development (along with institutional training and self-development). Implicit in this doctrine is the belief that Army leaders learn from their experience as leaders and that the lessons of job experience make a significant and independent contribution to leader development. An important objective of Army manpower and personnel research is, therefore, to explore the knowledge, skills, and abilities that contribute to effective leadership, as well as the on-the-job experiences that give rise to them.

The research described in this report was intended to further these organizational objectives. Specifically, it was intended to support the future development of assessment instruments based on tacit knowledge in the leadership domain and the military setting. In what follows, we present the theoretical and methodological background for the study to be reported. We briefly describe the results of a prior study upon which the current study seeks to build. We describe the goals and methods of the current study, organizing this description around three major data sets. We describe the analysis of these data sets and report results separately for each. Finally, we seek to fit results, drawn from all three data sets, into an emerging validity profile for the assessment instrument to be developed in the next phase of this project. In so doing, we hope both to integrate the results obtained from the three data sets as well as to set the stage for future instrument development.

Tacit Knowledge and the Lessons of Experience

A body of research in the behavioral sciences, with roots in both the information-processing and psychometric research traditions, holds promise for understanding the process of learning from experience in knowledge-intensive disciplines. In a sense, this research issues from a single, simple observation--that learning from experience often occurs without conscious intention to learn or conscious awareness of having learned. Rather, such learning is experienced as something that happens "behind the scenes" as people pursue goals on the job. The common language of the workplace reflects an awareness of this fact as people speak of "learning by doing" and "learning by osmosis." No less an observer of human mental life than William James remarked on the implicit quality of on-the-job learning in his discussion of "pedagogical implications" of the laws of habit (James, 1890).

> Let no youth have any anxiety about the upshot of his education, whatever the line of it may be. If he keep faithfully busy each hour of the working day, he may safely leave the final result to itself. He can with perfect certainty count on waking up some fine morning, to find himself one of the competent ones of his generation, in whatever pursuit he may have singled out. Silently, between all of the details of his business, the *power of judging* in all that class of matter will have

built itself up within him as a possession that will never pass away. (p. 127)

When learning occurs implicitly, behind the scenes, the knowledge that results has a tacit quality--people may be unaware of what they know and may have difficulty articulating it, even when prompted. Again, the language of the workplace is instructive. Terms such as "professional intuition" and "professional instinct" seem intended to denote the opaque or tacit quality of knowledge gained from job experience. In this section, we briefly describe research that supports the psychological reality and practical importance of tacit knowledge in professional competence.

The opaque quality of expert knowledge is, of course, well documented in the literature on human expertise (see Chi, Glaser, and Farr, 1988). Research with experts in a variety of knowledge-intensive domains has shown that reasoning and problem solving in such domains depend upon proceduralized skills and schematically-organized knowledge, both of which may operate outside of focal awareness. Further, expert knowledge may reflect the structure of the operating environment or situation more closely than it does the structure of formal, disciplinary knowledge (Groen & Patel, 1988)--making a focus on such formal knowledge a relative "blind alley" in efforts to understand expert performance. Experts queried about what they know often have great difficulty articulating the knowledge that underlies their decisions or capabilities on the job. Indeed, the great resistance of expert knowledge to articulation and codification has spawned extensive research on methods of elicitation--from structured interviews, to q-sort procedures, to repertory grid techniques. Despite this work, the opaque quality of expert knowledge continues to represent a major "bottleneck" in the development of expert systems and other intelligent computer-based applications.

Further support for the psychological reality of implicit learning and tacit knowledge comes from research, conducted in the laboratory, focusing on the phenomena of learning without intention or awareness. The foundational research in this area was conducted in the late 1960s by Arthur Reber and colleagues (Reber, 1967; Reber & Millward, 1968; Reber, 1969). Their work on the acquisition of stochastic grammars and of event sequences suggested that human subjects are capable of acquiring knowledge of a very complex nature without conscious intention or awareness of learning. Later researchers applied the paradigm to study learning of meaningful information (e.g., information about other people, information about the behavior of an economic system) and replicated the basic pattern of results (Broadbent & Aston, 1978; Broadbent, Fitzgerald, & Broadbent, 1986). Laboratory work on implicit learning suggests that subjects are able to exploit the structure inherent in a stimulus display in order to gain useful knowledge of the regularities in their environment. Importantly, this knowledge seems to be acquired in the absence of awareness or intention to learn--it is knowledge of a hidden or tacit nature.

Support for the practical relevance of tacit knowledge comes from research conducted in the world of professional practice (managerial, academic) focusing on

individual differences in tacit knowledge and on the consequences of these differences for professional success (Sternberg, Wagner & Okagaki, 1993; Wagner & Sternberg, 1985; Williams & Sternberg, in press). Because this research provided the major scientific justification for the work described in this report, we remark on its major findings. First, however, we must say a bit more about tacit knowledge as a theoretical construct.

The Tacit-Knowledge Construct

The example of battalion commander knowledge offered at the beginning of this paper serves to underscore the key features of tacit knowledge as a construct in our research. First, and most importantly, tacit knowledge is knowledge that is generally acquired on one's own--through personal experience rather than through instruction. Second, tacit knowledge is knowledge that people may not know they possess and/or may find it difficult to articulate. Like much expert knowledge, it is knowledge that guides behavior without being readily available to conscious introspection. Obviously, the hidden or opaque quality of tacit knowledge is the feature that gives the construct its name.[1] Finally, tacit knowledge is action-oriented knowledge with practical value to the individual. Unlike much disciplinary knowledge, it is knowledge that helps people pursue goals that they personally value.

Thus, as the construct has been employed in the research program of Sternberg and colleagues, tacit knowledge is a subset of all job-relevant knowledge, but a subset with special properties.[2] These special properties (acquisition on one's own, resistance to introspection, and practical value) make tacit knowledge particularly useful for understanding intelligent behavior in real-world settings, as well as for predicting success in such settings (to the extent that individual differences are present in the ability or inclination to acquire and use tacit knowledge).[3]

A second conceptualization of tacit knowledge, consistent with the featural model specified above, treats tacit knowledge as a cognitive phenomenon and defines it in terms of the learning processes that produce it and the memory structures/ systems that encode it. This model, which we shall refer to as the explanatory model of tacit knowledge, has several advantages over the featural model for our current

---

[1] The term "tacit knowledge" has roots in works on the philosophy of science (Polanyi, 1966), ecological psychology (Neisser, 1976), and organizational behavior (Schön, 1983). The adaptation of the term to account for individual differences in practical intelligence reflects an intellectual debt to all of these sources.

[2] The claim that tacit knowledge is a type of job knowledge is attributable to Schmidt & Hunter (1993).

[3] Note that the feature "practical value" is used here to capture the key content of the feature "procedural structure" used in some of our prior writings on tacit knowledge. The relationship between tacit knowledge and proceduralized skill is discussed in Appendix B.

purposes. Most importantly, it can be used to support the interpretation of scores on tacit-knowledge tests (the development of which the current study is intended to support) by embodying a high-level model of performance on such tests. That is, it provides a rationale for relating scores on tacit-knowledge tests to the underlying tacit knowledge of test takers. Note, however, that the extension of the explanatory model is essentially the same as that of the featural model--both models tend to "pick out" or point to the same knowledge. That is, both the featural and the explanatory model of tacit knowledge refer to knowledge that is acquired on one's own, not readily articulated, and relevant to action directed toward personally valued goals.

An explanatory model of tacit knowledge.

The explanatory model of tacit knowledge begins with a basic distinction between episodic and semantic of memory, attributable to Tulving (1972).[4] Episodic memory is defined as memory for specific, personally experienced events--memory for the "episodes" that make up one's experience. For example, an Army officer's memory of the weather, activities, and communications that made up his last NTC[5] rotation is classified as episodic. The hypothesized contents of episodic memory are often described as cases, situations, or event representations. Semantic memory is defined as memory for general, impersonal knowledge--memory for information that transcends particular episodes. For example, an officer's memory of the typical features of an NTC rotation (e.g., fast paced, disrupted communications, exhaustion) is classified as semantic because it is generalized knowledge and does not depend upon memory for a particular rotation (such as an officer's second, at which things may have gone smoothly). A major focus of research on memory and cognition during the last twenty years has been the transition from episodic storage to semantic storage of information about personally experienced events. That is, how does an officer's memory for several NTC rotations become assembled into knowledge of NTC rotations in general--knowledge that he can draw on long after his memory for the details of a particular rotation are lost?

According to models of inductive learning (e.g., Holland, Holyoak, Nisbett, & Thagard, 1986), the transition from event knowledge to generalized knowledge involves mental processes that are sensitive to the covariance structure of the environment, to "what goes with what" in the world. These processes (variously referred to as induction, abstraction, or extraction of invariants) isolate shared features and/or structure across episodes and construct abstract or general representations of that shared structure. Thus, for example, an officer's experience of a training exercise very frequently includes sleep deprivation and, on the basis of that commonality across episodes, his generalized knowledge representation includes

---

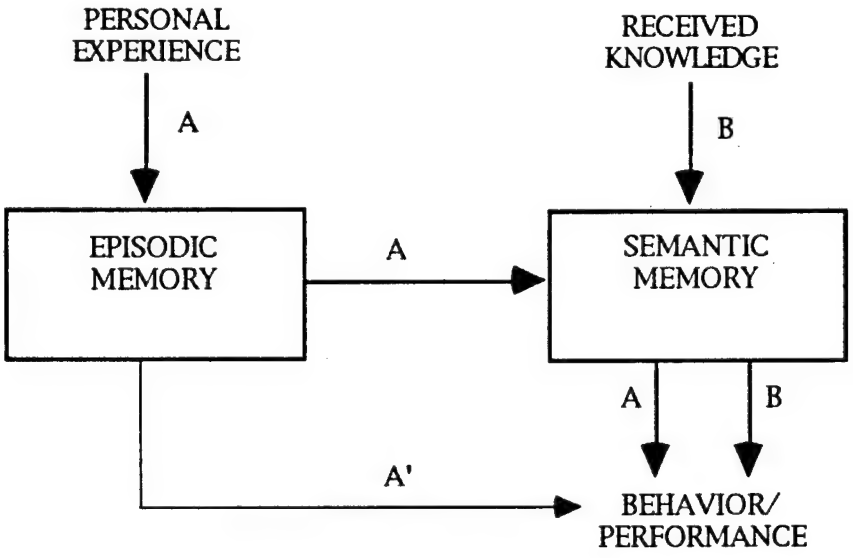
[4] In more recent work, Tulving has hypothesized a third memory store consisting of procedural knowledge. The relationship between tacit knowledge and procedural knowledge is discussed in Appendix B.

[5] The National Training Center (NTC) is an environment for high-fidelity combat simulation.

the feature "exhaustion" and leads him to expect that he will have to cope with exhaustion.

Figure 1 shows the two memory stores (episodic and semantic) along with an arrow between them that represents the process whereby high-level invariants are identified and stored as generalized knowledge structures in semantic memory. If we treat the top of the figure as representing the stimulus environment (the source of inputs to the memory system) and the bottom of the figure as representing behavioral consequences of learning (the output of the memory system) then we may use our very basic memory model to good effect in understanding the tacit-knowledge construct.

Figure 1.
Memory Structures and Knowledge-Acquisition Pathways in a Explanatory Model of Tacit Knowledge



First, note that we can identify three major pathways through the memory system. The first pathway, labeled A in the figure, corresponds to the process by which personally experienced events are stored in episodic memory and, over time, used to construct generalized knowledge structures in semantic memory. Thus, Path A can be seen as a pathway of experiential learning--the operation of memory structures and processes by which personal experiences become generalized knowledge of the world.

Path B, by contrast, corresponds to the process by which generalized knowledge of the world is acquired directly--most typically through a process of formal instruction. For example, a civilian researcher might have no personal experience at NTC yet still acquire knowledge of "typical" features of NTC rotations by reading about them in Army literature. Such knowledge, according to our model, takes the form of "received knowledge" that is input, more or less directly, to semantic memory. Like knowledge that enters semantic memory through Path A, Path B knowledge may provide a basis for intelligent behavior or performance. Thus, the civilian researcher may visit the NTC as an observer and know, based on his reading, that he should expect to work long hours.

Although both paths in Figure 1 may produce generalized knowledge that supports behavior/performance, Path B may be distinguished from Path A according to the processes by which that knowledge comes to be represented in semantic memory. Path B knowledge is knowledge that has been pre-processed for the learner (perhaps by an author, teacher, or advisor) whereas Path A knowledge is knowledge that the learner has processed on his or her own--through the accumulation of personal experiences, the extraction of high-level invariants from those experiences, and the construction of generalized knowledge structures based on those invariants. This distinction, between Path A knowledge and Path B knowledge, in terms of "who" did the processing (self vs. other), will be seen to have consequences for the relative usefulness of the resulting knowledge representations.

Consider a third path in Figure 1, labeled Path A'. This path represents the direct influence of event knowledge in episodic memory on behavior--influence that is not mediated by the generalized knowledge representations in semantic memory. The psychological reality of such direct influence has been demonstrated in a variety of experimental paradigms (e.g., Jacoby, 1983; Schacter, 1987). Even when knowledge processing has taken place, and memory for individual episodes appears to be lost, information about those episodes can continue to influence behavior.[6] Thus, Figure 1 shows three knowledge acquisition pathways; two (Paths A and A') represent learning from personally experienced events and one (Path B) represents the direct acquisition of generalized knowledge, typically through formal instruction.

Having contrasted learning from personal experience (and the construction of generalized knowledge structures based on that experience) with learning of received knowledge, it is a simple and obvious matter to define tacit knowledge as knowledge acquired via the former (Paths A or A') and to contrast tacit knowledge with job-relevant knowledge acquired via the latter (Path B). To do so is to identify mental processes and structures that result in knowledge that is (1) acquired on one's

---

6 Interestingly, when memory of personally experienced events influences behavior directly, it often takes the form of knowledge that people do not know they have and/or cannot readily articulate. The most direct support for this proposition is found in studies of implicit memory, a typical example of which is described in Appendix A.

own, (2) not readily articulated, and (3) relevant to action directed toward personally valued goals. Put another way, the key features of tacit knowledge can be seen to "load" very strongly on Paths A or A' and to load very weakly on Path B.

First, knowledge acquired via Paths A or A' is, by definition, knowledge acquired through personal experience. Second, knowledge that is acquired via Path A' is likely to be knowledge that is not readily articulated. Such knowledge reflects the direct influence of memory for particular episodes which, as we have said, often cannot be readily articulated. Finally, knowledge that is acquired via Paths A or A' is likely to be knowledge that supports action directed toward personally valued goals because such knowledge is acquired during the course of goal-directed action. Path B knowledge, as we have defined it, is not acquired through personal experience but through the communication of generalized knowledge based on someone else's experience. Because it has been formulated for communication, Path B knowledge is knowledge in a readily articulable form. Finally, Path B knowledge may vary in its relevance to personally valued goals to the extent that those goals differ from the goals of instruction. Thus, the featural description of tacit knowledge "falls out" of an explanatory model based on what we regard as an uncontroversial, if simplified, memory model.

The explanatory model also provides an account of how knowledge acquired via Paths A or A' (i.e., tacit knowledge) confers a performance advantage in certain classes of situations. That is, when "behavior/performance" in Figure 1 is defined as responses to complex, contextualized problem situations within the target domain, then a performance advantage should accrue to the individual who has acquired relevant tacit knowledge. We offer two basic arguments in support of this proposition. First, knowledge acquired via Path B will not support behavior/performance through the direct influence of stored cases/events.[7] As described above, such direct influence is often manifested tacitly–in the form of intuitions, or by facilitating the search for solution-relevant information in memory. Because knowledge acquired via Path B is not based on memory for episodes, such memory cannot be brought to bear, implicitly or otherwise, on performance.

A second argument for the primacy of tacit knowledge in responding to realistic, contextualized problem situations concerns differences in the quality and applicability of generalized knowledge acquired via Paths A and B, respectively. When generalized knowledge (e.g., a model of attitude change in work settings) is brought to bear on a realistic, contextualized problem situation, the likelihood of a fit between that knowledge and the features of the situation will be a complex function of the variability among and representativeness of the episodes on which the knowledge is based. To the extent that one's past experiences are predictive of one's future experiences, generalized knowledge acquired via Path A should have a higher

---

[7] It is unclear whether high-quality, case-based teaching bears a stronger resemblance to learning from personal experience or to learning received knowledge. We tend to favor the former.

probability of "fit" to the current situation, when compared to generalized knowledge acquired via Path B.[8]

Having presented an explanatory model of tacit knowledge, several *caveats* are in order. First, our purpose in proposing an explanatory model is to provide a basis for assessing and establishing the construct validity of tacit-knowledge tests. Thus, the foregoing argument sought to draw broadly-based distinctions between classes of mental processes, as well as between classes of resulting knowledge, in order to say what tacit knowledge is and is not. An unintended consequence of drawing such distinctions has been to ignore the overlap and interaction between the two knowledge-acquisition pathways. Clearly, generalized knowledge acquired via Path B is used to reflect on and clarify the meaning of both experienced events and high-level generalizations based on those events. Conversely, knowledge based on personally experienced events is used to instantiate, evaluate, and modify the generalized knowledge acquired via Path B. Thus, although the purpose of the foregoing discussion has been to distinguish tacit knowledge from other job-relevant knowledge, it would be naive to suggest that these two types of knowledge exist in any "pure" form.

A second unintended consequence of the distinctions drawn above is to appear to argue for the general superiority of knowledge acquired via Paths A or A'. However, the discussion of tacit knowledge as a cognitive process should not be taken to imply that tacit knowledge is superior to formal or disciplinary knowledge for all purposes. Clearly, for many purposes--explaining one's actions to others, reflecting on and judging the representativeness of one's experience, seeking to relate what one has learned to a wider body of disciplinary knowledge--knowledge acquired via Path B is superior to knowledge acquired via Paths A or A' (which does not support these activities well). Perhaps most importantly, knowledge acquired via Path B will be superior to that acquired through Paths A or A' whenever an individual's experience has been minimal and/or unrepresentative with respect to the problem or situation in terms of which behavior/performance is defined.

These caveats notwithstanding, the principle of transfer-appropriate processing predicts that, *ceteris paribus*, the more event-like the performance measure, the greater will be the contribution of event-based knowledge. That is, by extension, the more realistic and contextualized the situation to which a subject must respond, the greater will be the contribution to effective responding of tacit, experience-based knowledge. We will return to this idea in a later section when we discuss the interpretation of scores on to-be-constructed tacit-knowledge tests.

---

[8] By contrast, of course, when someone else's experience is more predictive of one's future experience than is one's own past experience (i.e., in situations of great novelty or at periods of career transition), then the level of fit between generalized knowledge and the current situation may be higher for knowledge acquired via Path B (as compared to Path A).

Empirical Research on Tacit Knowledge

Knowledge with the general properties outlined above has been studied in a program of research conducted by Sternberg and colleagues (for a recent review, see Sternberg, Wagner, Williams, & Horvath, 1995). This research has focused, primarily, on individual differences in the ability to acquire and use tacit knowledge, and on the consequences of those differences for performance in knowledge-intensive disciplines (e.g., academic psychology, business management, sales). In this section we briefly review the most relevant findings from this program of research.[9]

First, tacit knowledge can be effectively measured (Wagner, 1987; Wagner & Sternberg, 1985; Williams & Sternberg, in press). The measurement instruments employed in this research typically consisted of a set of work-related situations, each with between five and twenty response items. Each situation posed a problem for the subject to solve, and the subject indicated how he or she would solve the problem by rating the various response items. For example, in a hypothetical situation presented to a business manager, a subordinate whom the manager does not know well has come to him for advice on how to succeed in business. The manager is asked to rate each of several factors (usually on a 1 = low to 9 = high scale) according to its importance for succeeding in the company. Examples of factors might include (a) setting priorities that reflect the importance of each task, (b) trying always to work on what you are in the mood to do, and (c) doing routine tasks early in the day to make sure you get them done. The set of ratings the subject generates for all the work-related situations is the measure of his or her tacit knowledge for that domain. The procedure for scoring tacit-knowledge tests has undergone evolution across several studies, and a detailed description is beyond the scope of this article. In general, tacit-knowledge tests have been scored in one of three ways: (a) by correlating subjects' responses with an index of group membership (i.e., expert, intermediate, novice), (b) by computing the difference between subjects' responses and an expert prototype, or (c) by judging the degree to which subjects' responses conform to professional "rules of thumb."

Tacit knowledge has been found to increase, on average, with job experience, but it is not a direct function of job experience (Wagner, 1987; Wagner, Rashotte, & Sternberg, cited in Sternberg, Wagner, & Okagaki, 1993). What matters most is not how much experience a person has, but how well the person utilizes the experience to acquire and use tacit knowledge. As mentioned above, tacit knowledge is not a fancy proxy for IQ--at least within the range of abilities typical of those who normally enter managerial or professional occupations.[10] Tacit knowledge almost never correlates significantly with IQ. In the one case when an aspect of tacit knowledge did correlate significantly with IQ, that aspect was a particularly poor predictor of job performance (Wagner, Rashotte, & Sternberg, cited in Sternberg, Wagner, & Okagaki, 1993). Tacit knowledge also correlates trivially with other

---

[9] Portions of this discussion are taken from Horvath et al. (1994b).

[10] The relationship between IQ and tacit knowledge is undetermined in the general population.

conventionally measured abilities, in particular, those measured on the Armed Services Vocational Battery. Tacit knowledge is not a proxy for measures of personality, cognitive style, or interpersonal orientation. When tests of these attributes were given to managers, and hierarchical regression was used to predict performance on managerial simulations, tacit knowledge of management was the best single predictor of performance on the simulation (Wagner & Sternberg, cited in Sternberg, Wagner, & Okagaki, 1993). The contribution of tacit knowledge to prediction was still significant after holding all other variables constant.

Although tacit-knowledge measures do not correlate significantly with measures of potentially confounding constructs, subscores within a domain (e.g., tacit knowledge of self, others, or tasks) do correlate moderately with one another (about .3), suggesting that there may be a general factor underlying tacit knowledge, within a domain, that is different from the general factor measured by traditional psychometric tests of intelligence (Wagner, 1987; Wagner & Sternberg, 1985; Williams & Sternberg, in press). Tacit-knowledge scores also correlate across domains (at about the .5 to .6 level), suggesting that there is at least some commonality in the tacit knowledge required for success in different professions (Wagner, 1987). The tacit knowledge required for success in any setting has been found to depend upon the nature of the institution and the level of advancement one has reached within that institution.

Tacit knowledge predicts job performance moderately well, correlating about .3 to .5 with measures of rated prestige of business or institution, salary, performance appraisal ratings, number of publications, etc. (Wagner, 1987; Wagner & Sternberg, 1985; Wagner, Rashotte, & Sternberg, cited in Sternberg, Wagner, & Okagaki, 1993; Williams & Sternberg, in press). These correlations, uncorrected for attenuation or restriction of range, compare favorably with those obtained for IQ within the range of abilities we have tested. Tacit knowledge also predicts both academic performance and self-reported adjustment in a college setting (Williams & Sternberg, cited in Sternberg, Wagner, & Okagaki, 1993). Its prediction of the former is about as good as that of conventional academic-ability tests (with a multiple R of about .6), whereas its prediction of adjustment is better (with a multiple R of about .8).

To summarize, tacit knowledge can be measured, it increases with experience, and it can predict job performance, perhaps better than IQ, even when job experience is held constant. Tacit knowledge provides a significant increment of prediction above and beyond other psychological measures. Further, different aspects of tacit knowledge are correlated among themselves, suggesting the possibility of a general factor for tacit knowledge acquisition and use.

Summary

We have suggested that learning from experience on the job has an implicit or "behind the scenes" quality and that much of the knowledge acquired through job experience is of a hidden or tacit nature. We have summarized research, in a variety of traditions, that supports this general proposition. We have described a theoretical

construct, tacit knowledge, that holds promise for understanding the relationship between on-the-job learning and professional performance. The tacit-knowledge construct (previously defined in terms of characteristic features, now defined in terms of knowledge-acquisition pathways) has been shown to be relevant to and predictive of success or failure in knowledge-intensive disciplines. We have argued, on the strength of these findings, that the tacit-knowledge construct represents a promising avenue of inquiry in seeking to understand and promote the process of learning about leadership from experience. In the section to follow, we provide an overview of a multi-year research effort, of which the research described in this report is a part, that applies the tacit knowledge construct, and associated methods, to the problem of understanding and optimizing leader development through job assignments in the Army.

Overview of the Project

The research described in this report is part of a larger effort to apply the theory and methods of tacit-knowledge research to the problem of leader development in the Army. Specifically, the project seeks to identify (and provide instruments for measuring) the practical, action-oriented knowledge that military leaders acquire through job assignments. It further seeks to apply this knowledge to the problem of leader development through operational assignment. To these ends, the project is divided into three phases: (1) identification of tacit knowledge, (2) construction and validation of assessment instruments for measuring tacit knowledge, and (3) application of tacit knowledge and associated instruments to leader development.

The identification phase of the project, now completed, included a systematic review of the military practice literature (see Horvath et al., 1994a) and an empirical study of "leadership lessons" learned by Army leaders at the three levels under study (see Horvath et al., 1994b). The latter study was based on a series of semi-structured interviews with 81 Army officers and involved the elicitation, coding, and analysis of 179 items of leader knowledge meeting the criteria for tacitness stated above. The methods and results of this study are described in the following section of this report.

The assessment phase of the project, currently in progress, includes a study designed to support the development of tacit-knowledge tests (i.e., the study described in this report), the construction and validation of tacit-knowledge tests at each of the levels under study, and the construction of test manuals and supporting materials for the use of these tests. Finally, the application phase of this project will seek to evaluate the results of the earlier phases of work in terms of their applicability to the "real world" of leader development through operational experience. As currently envisioned, this phase of the project will focus on supporting and training those who have responsibility for developing officers at the specified levels (a so-called "training the trainers" approach).

Goals of the Current Study

The primary purpose of the current study is to support the development of assessment instruments based on (a) the theory and methods of tacit-knowledge research, and (b) the substantive knowledge acquired from Army officers during the identification phase of the project. Specifically, the study was intended to select that content, from the corpus of tacit knowledge obtained in the interview study, that was most promising with respect to the goals of the assessment phase. The term "promising" is here used to refer to those tacit-knowledge items (or groups of items) with the highest probability of yielding or contributing to tacit-knowledge test questions that, taken together, constitute a valid measure of the underlying, domain-relevant tacit-knowledge of respondents. Thus, we sought to select tacit-knowledge content for use in constructing the scenarios and response options of which tacit-knowledge test items will be composed. Such a selection process was necessary for several reasons.

First, the interview study yielded too many items of tacit knowledge to include in a tacit-knowledge test of reasonable length, given basic assumptions about the organizational context in which the test might be employed. Second, the interview study provided little basis for distinguishing between tacit-knowledge that is diagnostic or predictive of leadership experience and/or leadership effectiveness and tacit knowledge that is unrelated to these criteria. Third, the interview study provided only preliminary evidence regarding the distribution of tacit knowledge items across functional aspects of the leader's role. That is, the interview study provided an insufficient (or uncorroborated) basis for the construction of tacit-knowledge tests whose internal structure captures or reflects the internal structure of the hypothesized construct domain (the tacit-knowledge space). Finally, prior experience in the development of tacit-knowledge tests has shown that the process of constructing scenarios and response options of high quality is extremely resource intensive. As a consequence, this intermediate study was designed to narrow down the range of content from which tacit-knowledge test items would be constructed.[11]

Thus, the goal of the current study was to set the stage for the construction of tacit-knowledge tests, the interpretation of scores on which would meet established standards of validity. In addition to the selection of content, the current study supports instrument development by providing an interim statement of our current position on several theoretical and measurement issues. To this end we have outlined an explanatory model of tacit knowledge that, we propose, forms the basis of a high-level task model for interpreting scores on tacit knowledge tests and for assessing the validity of score interpretations. Also to this end, we have sought to fit the results of the current study into a unified validity framework in order to show how our current findings will support our long term goal of valid, useful tacit-knowledge tests.

---

11 The intermediate study also serves to provide an "in process" check on the prospects for developing externally valid tacit-knowledge tests.

Identifying Tacit Knowledge for Military Leadership

In this section we summarize the results of a prior study in which a body of leadership knowledge was elicited, in semi-structured interviews, from active-duty Army officers around the United States. This knowledge--culled, codified, and analyzed by a panel of military leadership experts--provided the "raw materials" for the present study in which the relationship between tacit knowledge and leadership experience/effectiveness was examined. Note that the goal of this section is to provide necessary background for the present study. A detailed description of underlying theory, methods, and results of the interview study is provided in Horvath et al. (1994b).

Subjects from whom we acquired tacit knowledge were U.S. Army officers on active duty. They were drawn from each of three branch categories within the Army (combat arms, combat support, and combat-service support) and from each of three organizational levels (platoon, company, and battalion leaders). The proportion of females in the sample (.09) was comparable to that in the population of U.S. Army officers (approximately .10). Subjects were selected by brigade and/or battalion commanders who were asked to provide a representative sample of officers at each of the specified levels.

Subjects were told that the interviewers were interested in lessons about leadership that are not written in books or taught in classes and that the goal of the interview was to identify specific examples of informal knowledge about leadership at the subject's current level. The interviewer emphasized to the subject that he was not interested in doctrine or theory--the "party line" on Army leadership--nor in purely technical knowledge (e.g., supply procedures, gunnery, etc.). The interviewer signaled the beginning of the interview by asking the subject to tell a story about an experience from which he or she learned something about leadership. Asking for stories was a way to get subjects talking and to direct them towards concrete experiences and away from leadership theory. The goal was not only that subjects should tell stories but that they should express, in their own words, the leadership lessons learned in the situations described. Guidelines for follow-up questions were developed by the research team prior to conducting the interviews.[12]

After the interviews had been conducted and written interview summaries compiled, the tacit knowledge contained in the interview summaries was identified and coded. Two members of the research team served as raters in the preliminary stages of this identification and coding process. They reexamined the interview summaries and sought to identify knowledge that qualified as tacit knowledge for military leadership according to three criteria (described below). Degree of

---

[12] Note that the instructions given to interview subjects--to recall and reflect on personally experienced events from which they learned something about leadership--makes direct contact with the tacit-knowledge construct as defined in the explanatory model. That is, we sought explicitly to exclude Path B knowledge such as that acquired in courses, doctrinal manuals, and self-study.

interrater agreement was assessed for 18 of the 81 interview summaries, consisting of 48 leadership stories. Out of a total of 48 stories evaluated, the two raters reached agreement on 35 stories, or 73%.

When the tacit knowledge within each story had been identified by consensus of the two raters, each interview story was annotated with a preliminary coding of the tacit knowledge it contained. That is, each piece of identified knowledge was expressed as a mapping between a set of antecedent conditions and a set of consequent actions. An example of a tacit-knowledge story and the item derived from it is shown in Table 1. As the example shows, each item of knowledge was represented by one or more antecedent condition or "IF" statements, by one or more consequent action or "THEN" statements, and by a brief explanation or "BECAUSE" statement. The logical operators "AND" and "OR" were used in the coding to signal relationships of conjunction and disjunction, respectively. The programming construct "ELSE" was employed in the coding to connect sets of condition-action mappings into more complex procedures.

Table 1.
Example Leadership Story With Coded Knowledge Item

Story Summary

The battalion commander noticed that his company commanders were trying so hard to be successful that they would accept missions that their units did not have the capabilities to execute. Thus, the companies and the commanders would expend a great deal of effort and time to accomplish the mission without asking for help from the battalion in order to demonstrate their talents as leaders. The battalion commander gave one of his commanders a mission and the commander worked his unit overtime for two weeks to accomplish it. The battalion commander realized that the same mission could have been accomplished in two days if the commander had requested resources from the battalion. After that incident, the battalion commander made it a point to ask the company commanders to realistically assess their units' resources before taking on a mission. The battalion commander felt that all commanders wanted to succeed and earn the top block rating due to the competitive environment in today's Army.

Coded Item

IF your company commanders have a strong desire to be successful and earn top block ratings
AND
IF they also have a tendency to take on resource-intensive missions that exceed their capabilities
AND
IF they are reluctant to ask higher headquarters for help when they have missions that tax their units' resources
THEN require commanders to conduct resource assessments before they take on missions
BECAUSE an accurate resource assessment should indicate whether or not the unit has the resources to handle the mission. This assessment may prevent commanders from taking on a mission that would overburden their unit.

When the two raters had completed the preliminary coding of the tacit knowledge contained in the interview data, the annotated summaries were routed to an expert panel consisting of the three senior military members of the research team. These individuals, who together possessed 72 years of military-leadership experience, independently evaluated the preliminary identification and coding of tacit knowledge. When the members of the expert panel had made amendments to the coding, they met as a group to discuss and reach consensus on the final coding of the tacit knowledge contained in the interview summaries. The result of this meeting was a set of 174 coded knowledge items, representing the expert consensus on the tacit-knowledge content of the interview data.

The obtained tacit-knowledge items were grouped into categories for purposes of data compression and qualitative analyses. Members of the expert panel

independently sorted the tacit-knowledge items into categories of their own devising. Each individual performed three sortings: one for battalion commanders' tacit knowledge, one for company commanders' tacit knowledge, and one for platoon leaders' tacit knowledge. Individuals were free to form categories of whatever size and according to whatever rules of inclusion they wished, the only requirement being that categories be nonoverlapping.

The results of the independent sortings were used to form a set of dissimilarity matrices (one for each level) which were then cluster analyzed using a joining algorithm. Cluster analysis is a family of techniques for uncovering the natural groupings in a set of data (for a comprehensive review see Hartigan, 1975). The joining algorithm produces hierarchically organized clusters of items in the form of a tree. The hierarchical trees that resulted from the cluster analyses were interpreted by members of the expert panel. That is, the high-level subclusters in each tree were labeled, and the labeled clusters were taken to represent categories of tacit knowledge. Aggregating the results of the independent sorts through cluster analysis allowed us to identify those items that were grouped together according to multiple sort criteria. Put another way, the method of independent sorting and cluster analysis provided a more adequate sample of the population of possible sort criteria and thus increased our confidence in the validity of the resulting category structures.

Table 2 shows high-level categories of tacit knowledge that emerged from independent sorting and hierarchical cluster analysis. The numerical values in the table show the proportion of items at each level (battalion, company, platoon) that made up a given category. For example, items from the category "Protecting the organization" made up 9 of the 67 total items obtained from battalion commanders, yielding a proportion of .13. A blank line in Table 2 means that the indicated category did not emerge from cluster analysis at the indicated level. For example, the category "Protecting the organization" emerged for battalion commanders but not for company commanders or platoon leaders. Asterisks in Table 2 indicate between-level differences in proportion that were marginally or fully significant by chi-square tests of association.

Table 2.
Categories of Tacit Knowledge With Proportion of Items Obtained, by Level

Category	Level		
	Battalion	Company	Platoon
Dealing with poor performers	.06	---	---
Managing organizational change	.04	---	---
Protecting the organization	.13	---	---
Balancing mission and troops	---	.08	---
Cooperating with others	---	.06	---
Directing and supervising subordinates	---	.16	---
Establishing credibility	---	---	.12
Developing subordinates	.18**	.06	---
Influencing the boss	---	.08	.14
Communicating	.15	.13	.13
Establishing trust	.07	.08	.07
Managing the self	.07	.09	.19*
Motivating subordinates	.09**	.14	.28*
Taking care of soldiers	.14	.12	.05
Unaffiliated items	.07	0	.02

* $p = .07$, ** $p < .05$,

To summarize, our analysis by category of tacit knowledge for military leadership indicated a number of areas of tacit knowledge common to all three levels under study. These included tacit knowledge for "Communicating," "Establishing trust," "Managing the self," "Motivating subordinates," and "Taking care of soldiers." These categories may represent level-invariant areas of leader knowledge acquired through experience. Several level-specific categories also emerged, however, and these salient or distinguishing categories were the focus of analysis and discussion during the identification phase of the project. Salient categories at a given level

were found to reflect the developmental challenges created by changing role requirements that leaders typically face at that level. As such, they provided potentially valuable evidence concerning the changing focus of on-the-job learning as leaders ascend the organizational hierarchy. A detailed discussion of these findings may be found in Horvath et al. (1994b).

Methods

As described earlier, the current study was designed to support the construction of tacit-knowledge tests for use in U.S. Army leader development. The tacit-knowledge items acquired in the interview study formed the raw materials for this construction process. We presented the tacit-knowledge items to Army officers in three institutional settings, obtained rating and sort data from these officers, and analyzed the data in a variety of ways. In collecting and analyzing data, we sought to answer two basic questions about the tacit-knowledge sample. First, which items are most promising for use in the construction of tacit-knowledge test questions? Second, what does the latent structure in the tacit-knowledge corpus tell us about the structure of the construct domain and, in turn, about how we should structure our tacit-knowledge tests? In this section, we describe the methods by which the 174 tacit-knowledge items acquired in the interview study were presented to a larger sample of Army officers.[13]

Relationship of Tacit Knowledge to Experience (TRADOC)

The U.S. Army Training and Doctrine Command (TRADOC) conducts continuing professional education for active-duty Army officers at Army schools around the country. During his or her career, an Army officer will cycle between operational assignments and enrollment in TRADOC schools. At lower leadership levels (from pre-platoon through company level) TRADOC schooling is branch specific. That is, officers attend a school run by their branch (e.g., infantry, signal, quartermaster) that prepares them for leadership/command at the next highest level. Thus, for example, an infantry captain who has completed platoon leadership and who is "in the running" to command a company will be sent to the Officer's Advanced Course to be trained in knowledge and skills needed to command an infantry company and to serve as a battalion staff officer. At command levels above that of the company, TRADOC schooling is integrated across branches, consistent with the need to prepare senior leaders for cooperation with and command of combined-arms units.

TRADOC schools were selected as a source of subjects for two reasons. First, TRADOC schools provide a ready pool of active-duty officers at the levels under study. Second, the nature of selection to each school allowed us to designate respondents as experienced or novice leaders at a given level. By administering to these subjects a survey asking for ratings of tacit-knowledge items, we hoped to explore the relationship between leadership experience and ratings of tacit knowledge at a given level.

---

13 Larger than the sample from which the knowledge items were acquired.

Tacit-knowledge survey data were collected, through the mail, from thirteen TRADOC schools in the continental United States. These included Officer's Basic and Advanced courses for officers in the quartermaster, infantry, signal, combined logistics, engineer, and field artillery branches. TRADOC sites also included the Basic course for officers in the transportation branch, the Command and General Staff College, and the Army War College. We obtained random samples of students in each of these schools/courses. Sampling was conducted by a member of the research team using class rosters provided by points-of-contact at the TRADOC schools.[14] Once students had been selected for participation in the study, a packet was assembled for each subject and packets for each school/course were shipped to points-of-contact at each site. Each packet contained a cover letter from a senior member of the research team (Colonel and USMA[15] professor). The cover letter explained the survey and asked subjects for their support in the data-collection effort. Each packet contained one copy of a Tacit-Knowledge Survey (TKS) appropriate to the assigned condition (Level X Experienced vs. Novice) of the subject in question. The front matter of each TKS included explanatory text relevant to informed consent. Packets were distributed to subjects who completed the enclosed TKS on their own time and returned them to their point-of-contact. When all surveys had been returned, the point-of-contact bundled and shipped the completed surveys to Yale where civilian researchers secured them, inventoried them, and entered the TKS data into a computerized data base by means of an automated data-entry program.

Each respondent in the TRADOC sample was designated as either an "experienced" or a "novice" practitioner of leadership at one of the three levels under study. The respondent's status as an experienced or novice leader was determined by the course in which the respondent was enrolled and the respondent's previous command experience. Thus, respondents enrolled in the Officer's Basic Courses (Quartermaster, Infantry, Transportation, Signal, Engineer, and Field Artillery) were designated as novice platoon leaders because students in these courses have not yet led platoons (at least as officers). Respondents enrolled in the Officer's Advanced Courses (Infantry, Signal, Combined Logistics, Engineer, and Field Artillery) were designated as either novice company commanders (because students in these courses have not yet held command) or as experienced platoon leaders (because students in these courses have all led platoons and met with success sufficient for selection to the Advanced Course for their branch). Note that the designation of Advanced-Course students as either experienced platoon leaders or novice company commanders was made at random. That is, for these subjects a designation of either "novice company commander" or "experienced platoon leader" amounted to a random assignment to a category and a survey form. Thus, an Advanced Course student designated as a novice and an Advanced Course student designated as experienced differed only in the survey they were assigned to complete. Respondents enrolled in the Command and General Staff College were

---

[14] In the case of the Army War College, a point-of-contact conducted the sampling on-site.

[15] U.S. Military Academy at West Point, NY.

designated as either experienced company commanders or as novice battalion commanders (again, at random). Finally, students at the Army War College who participated in the study were designated as experienced battalion commanders (all had previously commanded a battalion).

Tacit-knowledge survey (TKS).

The tacit-knowledge surveys (TKS) were the basic vehicles for presenting tacit knowledge obtained in the interview study to a larger sample of Army leaders for evaluation. There was a single TKS constructed at each of the levels under study (battalion, company, and platoon) consisting of 66, 67, and 46 items, respectively. Each item presented a piece of "leadership advice" based on an item of tacit knowledge obtained from leaders in the interview study and at the specified level. That is, the knowledge items in each TKS were obtained from Army officers describing their experiences at that level of command. Subjects were asked to rate each item on each of four dimensions, as described below.

The development of tacit-knowledge surveys took place in three phases: planning, item rewriting, and scale construction. In the planning phase, consideration was given to the level of complexity at which tacit-knowledge items should be presented to subjects. Sample items at varying levels of complexity and with varying structures were generated and routed among project participants. An algorithm was developed for the simplification of tacit-knowledge items. This algorithm employed basic operations such as the condensation, abstraction, and deletion of information that made up tacit-knowledge items.

In the item-rewriting phase, the simplification algorithm was used to rewrite all of the tacit-knowledge items. The rewritten items were then routed to project participants at USMA who made additional revisions in order to increase the comprehensibility of the items for a military audience and to preserve the intention of the interviewees who provided the knowledge items. The completed set of rewritten items was divided by level (battalion, company, platoon). In the scale-construction phase, the dimensions on which subjects were to rate tacit-knowledge items were determined.

In its final form, the TKS asked subjects to rate each piece of leadership advice (i.e., each item of tacit knowledge) on each of four seven-point scales designed to elicit the following judgments: (1) how good does the respondent think the advice is, (2) how commonly known does the respondent think the advice is, (3) how often, in the judgment of the respondent, do leaders at the specified level face situations such as the one described, and (4) to what extent does the advice match the respondent's personal concept of leadership?[16] Each of the scales in the TKS was intended to provide a different sort of information about the tacit-knowledge item being rated. The good scale was intended to assess the overall quality of the knowledge being

---

16 In reporting of results, these rating scales are referred to as the "good", "known", "often", and "concept" scales respectively.

rated. The known scale was intended to assess one possible index of tacitness (i.e., on the theory that knowledge whose acquisition is not well supported by the environment may be less commonly known than other knowledge). The often scale was intended to assess the generalizability or applicability across Army leadership settings of knowledge items. Finally, the concept scale was intended to assess subjects' implicit theories of leadership practice. Together, the four rating scales in the TKS were intended to provide a comprehensive but nonredundant picture of each tacit knowledge item for the purpose of evaluating each item's potential for development into tacit-knowledge test questions. An example question (tacit-knowledge item plus four rating scales) is shown in Figure 2. Complete Tacit-Knowledge Surveys, for all three of the levels under study, are available upon request.

Figure 2.
Example Question From Tacit Knowledge Survey (Company Commander Level)

If a training event scheduled by your battalion commander (e.g., ADA Battalion) conflicts with a training event scheduled by your supported-unit commander (e.g., Infantry Brigade Commander) and if the event scheduled by the supported-unit commander has potentially greater training value, then take a risk and give priority to the supported-unit commander's training event. By taking a risk to provide your soldiers with the best training, you earn their trust.

1. How good is this advice for company commanders?

1	2	3	4	5	6	7
Extremely bad			Neither bad nor good			Extremely good

2. How commonly known is this advice among company commanders?

1	2	3	4	5	6	7
Known by almost none			Known by some			Known by almost all

3. How often do company commanders face situations like this?

1	2	3	4	5	6	7
Almost never			Sometimes			Almost all the time

4. To what extent does this advice match your concept of leadership?

1	2	3	4	5	6	7
Does not match my concept of leadership at all			Matches my concept of leadership somewhat			Matches my concept of leadership very closely

Broad import of the TRADOC data.

The TRADOC data provided information about the relationship between endorsements or judgments of tacit-knowledge items and level of leadership experience. Put another way, we looked across the range of experience within a given level and sought to find effects in the rating data that might be related to differences in experience. Specifically, we wanted to see if ratings on items or groups of items distinguished or discriminated between experienced and novice leaders at a given level. Any such "discriminating" items would be of obvious value, given our focus on supporting the construction of tacit-knowledge tests that measure underlying knowledge gained from experience.[17] That is, because knowledge is presumed to be acquired through experience, knowledge items that discriminate between experienced and novice leaders may reflect the possession of experience-based knowledge and, as a consequence, possess strong face validity with respect to the tacit-knowledge construct. Further, discriminating items may hold promise for development into tacit-knowledge test questions that discriminate among leaders on the basis of acquired tacit knowledge.

Relationship of Tacit Knowledge to Effectiveness (FORSCOM)

The U.S. Army Forces Command (FORSCOM) is comprised of TO&E[18] units that carry out military missions. The Army units under this command are stationed, during peace time, at posts in the continental United States. Much of an Army officer's early career is spent in operational assignments in TO&E units. In such assignments an officer is part of a chain-of-command through which communication, command, and control are implemented. Thus, an officer assigned to a FORSCOM unit (e.g., a rifle company in an infantry battalion of the 82nd Airborne Division at Ft. Bragg) works in concert with subordinates (those officers or soldiers below him in the chain-of-command), peers (those in comparable or lateral positions within the chain-of-command), and superiors (those above him in the chain-of-command). Thus, for example, a rifle company commander has subordinates (the lieutenants who lead the platoons that make up the rifle company), peers (the captains who command the other companies in the infantry battalion), and a superior (the lieutenant colonel who commands the battalion).

TO&E units in FORSCOM were selected as a source of subjects for two reasons. First, because they constitute the war-fighting component of the U.S. Army, they are

---

17 Clearly, differences in item ratings (or patterns of item ratings) between experienced and novice leaders may be attributable to a number of factors (i.e., those that covary with experience). Perhaps most importantly, level of experience within a military organization is linked to retention and promotion which are themselves sensitive to ability/performance differences. Thus, strong arguments regarding the role of experience-based learning in any observed between-group differences cannot be made on the basis of the TRADOC data.

18 Table of Organization and Equipment (TO&E) is used here to denote units for which personnel and equipment have been specifically authorized.

the units in which leadership impacts most directly on the core mission of the organization. In other words, FORSCOM units possess high "face validity" with respect to military leadership. Second, FORSCOM units provide a large sample of incumbent leaders--leaders currently in the job--at each of the levels under study. Finally, FORSCOM units display an "intact" chain-of-command, consisting of Army leaders who work together closely, and in which subordinates, peers, and superiors of each officer can be identified for rating purposes.

In FORSCOM units we collected data from all the available members of the officer chain-of-command in each of approximately 30 battalions. Data collected were responses to the tacit-knowledge survey (TKS) and responses to a leadership effectiveness survey (described below) that asked subjects to rate the overall leadership effectiveness of the officers in their chain-of-command. Because the students in TRADOC schools do not work within an operational chain-of-command, it was not feasible to collect effectiveness ratings from the superiors, peers, and subordinates of officers assigned to the TRADOC schools. The goal of the FORSCOM data collection was to obtain information about the relationship between tacit-knowledge ratings and leader effectiveness. That is, because experience-based tacit knowledge is presumed to support effective performance, knowledge items that are highly correlated with rated leader effectiveness would possess strong face validity with respect to the tacit-knowledge construct. Further, such items may hold promise for development into tacit-knowledge test questions that converge with exogenous measures of leader effectiveness.

Both the content of the tacit-knowledge survey and the content of the leadership-effectiveness survey were specific to respondents' positions within the chain-of-command. That is, incumbent company commanders filled out a "Company Commander's Leadership Survey" (TKS) and a "Company Commander's Leadership Effectiveness Survey" (LES). In the former, they rated items of leadership advice on four scales (as in the TRADOC data collection). In the latter they applied global effectiveness ratings to their battalion commanders ("superior" rating), to themselves ("self" rating), to the other company commanders in their battalion ("peer" ratings), and to the platoon leaders in their company ("subordinate" ratings). The procedure was identical at the platoon and battalion levels, with two exceptions. First, platoon leaders did not rate their subordinates (usually non-commissioned officers). Second, battalion commanders did not rate their peers (who are often geographically remote from them). Commanders at the brigade level rated battalion commanders.

Leadership effectiveness survey (LES).

The leadership effectiveness surveys (LES) were the basic vehicles for obtaining measures of the leadership effectiveness of FORSCOM subjects. There was a single LES constructed at each of the levels under study (battalion, company, and platoon). Each question asked the respondent to rate the overall leadership

effectiveness of the officers in his or her command.[19] Effectiveness ratings were to be made with reference to all the leaders, at the specified level, that the respondent had known in his or her Army career. In order to prevent what was seen as likely "inflation" in effectiveness ratings, the following text was included in the front matter of the LES (in this case, for a brigade commander rating battalion commanders).

> On the following page we ask you for a global rating of the leadership of your battalion commanders. We are NOT interested in an assessment of their future potential, as would be reflected in the senior rater profile on the OER.[20] Instead, we want your honest rating of how good their current leadership is from your point of view. We realize that most officers in today's Army are good leaders. However, for this study, we need for you to distinguish degrees of "goodness" among quality officers. These ratings will be used for research only and they are confidential.

A sample question from the LES is shown in Figure 3. The full set of surveys (one for each of the levels under study) is available upon request.

---

[19] The term "respondent" is introduced here in order to distinguish between the individual who fills out a survey (the "respondent" in this report) and the individual about whom information is sought or on whom measurements are taken (the "subject" in this report). For example, when a platoon leader rates his company commander on the LES, the platoon leader is the respondent (but is not the subject) and the company commander is the subject (but is not the respondent). In the FORSCOM data, not all respondents were subjects in the study. Further, respondents were not the subjects of the LES they completed (with the exception of "self" ratings).

[20] Officer Evaluation Report (OER) is a formal performance appraisal.

Figure 3.
Example Questions From Leadership Effectiveness Survey (Brigade Commander Rates Battalion Commander)

Unit Code: ___ ___ ___ ___ ___ ___ ___

Leadership Effectiveness Ratings

Battalion Commander Ratings

Instructions: Think about the battalion commanders listed below who are under your command. Compared to all other battalion commanders you have known,

How good (effective) is the leadership of each battalion commander? Please circle the number under the statement that best corresponds to your rating for each battalion designation.

	The Best	One of the Best	Better than Most	As Good as Most	Not Quite as Good as Most but still gets the job done	Well Below Most	The Worst
	1	2	3	4	5	6	7
_____	1	2	3	4	5	6	7
_____	1	2	3	4	5	6	7
_____	1	2	3	4	5	6	7

Given the time consuming nature of the tacit-knowledge survey, and the somewhat sensitive nature of the LES, we devoted considerable attention to the manner in which these data were collected. In order to minimize time taken from the duty day of the officers involved, copies of the tacit-knowledge survey were mailed to a contact person in each battalion with the request that the surveys be distributed to and completed by officers in the chain-of-command prior to the day of the Yale/USMA visit. When subjects arrived with completed tacit-knowledge surveys in hand, as they did on all but a few occasions, the time required for subjects to be briefed on and complete the leadership-effectiveness surveys was less than 20 minutes. Leadership-effectiveness ratings of battalion commanders were also collected, in separate sessions, from their brigade commanders.

In order to allow the assignment of leadership-effectiveness ratings to the data records of the job incumbents being rated, each position within the chain-of-command was assigned a seven-digit code. During collection of the leadership-effectiveness data, subjects applied this code to the front page of both the leadership-effectiveness survey and the tacit-knowledge survey. After data collection in a given battalion was complete, members of the research team applied the appropriate seven-digit code to the ratings contained in each leadership-effectiveness survey (i.e., the code corresponding to the incumbent being rated) before sorting and bundling the data from that battalion. In addition to making possible the efficient management and analysis of leadership effectiveness data, the use of incumbent codes helped to assure subjects that confidentiality and personal anonymity would be protected. A civilian member of the research team offered verbal assurances to this effect at the beginning of every data-collection session. Every precaution was taken in the handling, storage, and mailing of the leadership-effectiveness data to ensure that commitments regarding confidentiality and anonymity were met.

Broad import of the FORSCOM data.

The FORSCOM data provide information about the relationship between endorsements or judgments of tacit-knowledge items and rated leadership effectiveness. That is, we looked within levels of leadership experience at each command level and sought to find effects in the rating data that might be related to differences in effectiveness. Again, we wanted to see if ratings on items or groups of items distinguished between effective and less effective leaders. Any such items would be of obvious value in future construction of tacit-knowledge tests.

Organization of Tacit Knowledge (USMA)

The U.S. Military Academy at West Point is the crucible in which many of the Army's most distinguished leaders have begun their development. The academy is staffed by a combination of civilian and military personnel. Among the military personnel are members of the academic administration, senior members of the faculty of the various academic departments, junior members of the faculty of the various academic departments, commissioned officers, noncommissioned officers and enlisted soldiers in various support roles, and the students or "cadets" who receive their undergraduate and military education at the academy. The Department of Behavioral and Sciences and Leadership at West Point was a source of subjects in the current study. We used this department as a source of subjects for two reasons. First, the sort task that these subjects performed required that they possess knowledge and personal experience of Army leadership and faculty and graduates of the Department of Behavioral Sciences and Leadership satisfy these criteria. Second, several investigators in the current project were members of the department and had access to colleagues and students willing to serve as subjects.

Fourteen USMA subjects (three colonels, six captains, and five second lieutenants) independently sorted the tacit-knowledge items obtained in the

interview study into categories of their own devising.[21] Each individual in the sample performed three sortings: one for battalion commanders' tacit knowledge, one for company commanders' tacit knowledge, and one for platoon leaders' tacit knowledge. Individuals were free to form categories of whatever size and according to whatever rules of inclusion they wished. The only requirement was that the categories be nonoverlapping. The results of the independent sortings were used to form a set of dissimilarity matrices (one for each level). Each dissimilarity matrix is a cases-by-cases or symmetrical data matrix in which cases are individual tacit-knowledge items and values in the matrix are integers representing the number of times a given pair of items was sorted into a single category. The dissimilarity matrices were the input to a series of scaling analyses designed to uncover latent structure in the sorting data. Thus, there was a single scaling analysis performed at each of the three levels under study. Note that each such analysis was based on all 14 subjects' sorting of all of the tacit-knowledge items at one of the levels. Details and results of these analyses are described in a later section.[22]

Broad import of the USMA data.

The goal of the sorting task was to obtain data that would support inferences regarding the way in which leadership tacit knowledge may be organized in the minds of the subjects who performed the sorts. By making such inferences, we hoped to obtain preliminary information about the internal structure of the construct domain as represented by the items obtained in the interview study. By deriving dimensions and Euclidean distances that best capture the similarity relations inherent in our subjects' sorting decisions, we hoped to obtain a working representation of the "knowledge space" for tacit, leadership knowledge. As described in an earlier section of this report, such a knowledge space will be useful in the future construction of tacit-knowledge tests. That is, a multidimensional representation of the tacit-knowledge space (even a tentative one) will help us to build tacit-knowledge tests whose internal structure fits that of the construct domain. If we have an idea of the dimensions that structure the mental representation of leadership tacit knowledge, we will be better able to build a set of tacit-knowledge test questions that "cover the territory" of leadership tacit knowledge, as best that territory can be determined at this time.[23]

---

[21] Because early data from the sorting task were obtained prior to the development of the TKS, subjects sorted the original (i.e., uncondensed) versions of the tacit-knowledge obtained in the interview study. Thus, the knowledge items rated in the TKS differ in format, though not in content, from the items sorted by subjects at USMA.

[22] Sorting data for four of the present subjects were also used in hierarchical cluster analyses in the interview study report (Horvath et al., 1994b).

[23] The structural aspect of construct validity is typically, and most adequately, addressed through an iterative process of test construction and validation.

Results

In this section we report the results of the current study, treating each of the three data sets (TRADOC, FORSCOM, USMA) separately. In the next section, we seek to integrate the three sets of results by presenting a unified framework for assessing the validity of our to-be-constructed tacit-knowledge tests. By fitting results from the three data sets into this framework, we hope to set the stage for the next phase of our research project.

Relationship of Tacit Knowledge to Experience (TRADOC).

Summary statistics.

Sample sizes and yields from the TRADOC data collection are shown in Table 3. As this table shows, yields were comparable across levels and conditions. No data were collected on characteristics of nonrespondents.

Table 3.
Sample Sizes and Yields (in parentheses) from TRADOC Sample

Level	Condition	
	Novice	Experienced
Battalion	102 (.68)	77 (.77)
Company	112 (.75)	115 (.77)
Platoon	260 (.87)	125 (.83)

Summary statistics for ratings on the TKS are shown in Table 4. The columns in this table represent the three forms of the TKS, one for each of the levels under study. Mean values in this table are grand means computed across questions and subjects. Standard deviations reflect the dispersion among question means. The mean values in Table 4 are based on 66, 67, and 46 question means for battalion, company, and platoon-level TKS, respectively. Between-level differences on the good scale were significant in a Kruskal-Wallis nonparametric test on ranks ($X^2(2)=7.81$, $p=.02$). The source of mean differences is indeterminate, however, as mean goodness ratings at the three levels were based on judgments by different subjects of different knowledge items.

Table 4.
Means and Standard Deviations (in parentheses) on TKS Rating Scales, by Level and Group (Experienced vs. Novice) in TRADOC Sample.

Rating Scale/ Group	Level		
	Battalion	Company	Platoon
Good			
Experienced	5.53 (1.57)	5.19 (1.82)	5.04 (1.76)
Novice	5.55 (1.54)	5.12 (1.75)	5.09 (1.64)
Known			
Experienced	5.04 (1.30)	4.58 (1.32)	4.39 (1.26)
Novice	4.95 (1.33)	4.76 (1.49)	4.52 (1.30)
Often			
Experienced	4.74 (1.61)	4.82 (1.51)	4.76 (1.38)
Novice	4.86 (1.44)	4.75 (1.49)	4.84 (1.24)
Concept			
Experienced	5.34 (1.73)	5.03 (1.92)	4.83 (1.88)
Novice	5.40 (1.66)	4.99 (1.82)	4.96 (1.72)

Rating-scale intercorrelations.

The intercorrelations among rating scales in the TKS are shown in Table 5. This table shows values of Pearson's *r* for intercorrelations of the four rating scales of the TKS. Correlation coefficients were computed on question means in order to remove the effects of within-subject response dependencies. The number of values contributing to each correlation coefficient was 66, 67, and 46 for battalion, company, and platoon-level TKS, respectively. All correlations were highly significant (p=.0001 for all coefficients but one, p=.003 for the remaining coefficient).

Table 5.
Intercorrelations Among TKS Rating Scales, by Level and Group (Experienced and Novice) in TRADOC Sample.

	Mean Good	Mean Known	Mean Often	Mean Concept
Mean Good				
Battalion		*.55*	*.41*	*.90*
Company		*.50*	*.45*	*.92*
Platoon		*.48*	*.41*	*.90*
Mean Known				
Battalion	.42		*.28*	*.55*
Company	.59		*.43*	*.51*
Platoon	.53		*.35*	*.50*
Mean Often				
Battalion	.41	.24		*.44*
Company	.54	.56		*.48*
Platoon	.44	.43		*.45*
Mean Concept				
Battalion	.90	.42	.44	
Company	.91	.60	.56	
Platoon	.85	.54	.47	

Note.
Coefficients for Experienced group shown in upper half-matrix and in italics.
Coefficients for Novice group shown in lower half-matrix and in plain text.
$p<.003$ in all cases.

A principal components analysis was conducted on each of the correlation matrices in Table 5 (one per level). At each level, the analysis yielded only one component with an eigenvalue greater than or equal to one, indicating a single, general factor that might be called "quality." Rather than forming a composite variable based on a weighted combination of the four scales, we decided to focus on the goodness scales in later analyses. Our reasoning was as follows. The "goodness" and "concept" scales were of a type that permit relatively straightforward inferences regarding the leadership knowledge of the respondent. That is, if a respondent tells us that she thinks a knowledge item is very good, or that it matches her concept of leadership very closely, then we may infer that she possesses knowledge or beliefs that are consistent with or supportive of the knowledge being rated. By contrast, responses on the "known" and "often" scale do not support straightforward inference regarding the leadership knowledge of respondents. That is, if a respondent tells us that she thinks a knowledge item is well known among company commanders, then we may infer only that she possesses this belief (i.e., a belief about what other leaders know) but may infer little about her own leadership knowledge. On these

grounds, we excluded both the "known" and "often" scales from inclusion in our composite dependent measure. Because the correlation between the remaining "good" and "concept" scales was nearly unity for all three levels, forming a weighted composite would offer no advantage over using one or the other of the scales exclusively. For this reason, we focus on the "good" scale in most of the analyses to follow.

Discriminant analyses.

We conducted a series of discriminant analyses on goodness ratings in the TRADOC data. Discriminant analysis is a technique for exploring the relationship between a single, categorical variable (groups) and a set of continuous variables (discriminating variables). A weighted, linear combination of the discriminating variables is derived, under a least-squares model, that maximizes the divergence between groups. The linear combination of discriminating variables (known as a canonical discriminant function or CDF) can be tested for significance and the correlations between discriminating variables and the output of the CDF (known as structure coefficients) provide indices of the discriminating power of variables.

In the analyses described below, cases were subjects, discriminating variables were goodness ratings on tacit-knowledge items in the TKS, and groups were the two levels of the experienced versus novice variable. The discriminant analyses enabled us to (a) assess the overall discriminating power, with respect to the experienced/novice variable, of goodness ratings on the combined set of tacit-knowledge items (through a significance test on the CDF), and (b) identify those tacit-knowledge items with the highest degree of discriminating power (by examining the structure coefficients).

A canonical discriminant function was computed that distinguished between experienced and novice groups for each of the levels under study. At each level, the canonical correlation coefficient based on the discriminant function was significant ($R=.73$, $p=.0006$; $R=.72$, $p=.0001$; and $R=.55$, $p=.0001$, for battalion, company, and platoon-level data respectively).[24] Thus, novice and experienced leaders responded differently to the tacit-knowledge items in the TKS, and this difference in overall pattern of responding is not likely due to chance. We take this result as evidence of the promisingness of the tacit-knowledge corpus with respect to our eventual goal of developing valid tacit-knowledge tests. We reason that if variance due to between-group differences in underlying, experience-based knowledge contributed to the

[24] The canonical discriminant function was tested for significance using the canonical correlation between the output of the function and a categorical variable representing group membership. Probability values are for the null hypothesis that the correlation coefficient was drawn from a population with expected value zero.

discriminating power of the CDF, then tacit-knowledge tests based on items in the TKS may be similarly sensitive to differences in underlying knowledge.[25]

Of course, not all of the knowledge items in the TKS contributed equally to the discrimination between experienced and novice leaders. Thus, an important goal of this analysis (and of the study as a whole) was to identify those knowledge items with greatest promise for development into tacit-knowledge test questions. In the context of the TRADOC data, this meant identifying those items that most strongly discriminated between experienced and novice groups. For this purpose, we computed structure coefficients for each of the items in the TKS. These coefficients, (analogous to factor loadings in factor analysis) represent the correlation between values of a variable and the output of the canonical discriminant function. The higher the absolute value of a structure coefficient for a given variable the more "discriminating" that variable was between groups. Thus, we used the structure coefficients for the tacit-knowledge items in the TKS to identify those items with the highest discriminating power and, by extension, the greatest promise for development into tacit-knowledge test questions.

To reiterate, the value of the structure coefficient for a given item tells us how goodness ratings on that item were associated with overall patterns of responding that were either experienced-like or novice-like. Items with the highest and lowest structure coefficient values would be judged (on this criterion) to be the most promising items for incorporation into tacit-knowledge test questions.[26] Note that this selection criterion provides an advantage over a simpler, point-biserial correlation between item rating and group. The latter tells us only about the content of the item in question. The former tells us about the underlying complex of either experienced-typical or novice-typical responses and about a given item's participation in that complex.

"Knownness" ratings and discrimination.

In order to shed light on the tacit-knowledge construct we examined the relationship between ratings on the "known" scale and the discriminating power of items. We reasoned as follows. If knowledge in our sample is indeed tacit (under the construct definition put forth earlier) then it is knowledge that subjects acquired on their own and, as a consequence, may be less commonly known than knowledge of a nontacit nature (i.e., received knowledge). Further, if our canonical discriminant function has picked up on between-group differences in experienced-based learning, then the most tacit items in our sample may be the items with the greatest

---

[25] The relationship between ratings of tacit-knowledge items and the form of hypothesized, underlying leader knowledge raises some non-trivial measurement issues. We attempt to address these issues in the final section of this report.

[26] Because it is a correlation coefficient, the magnitude and sign of a structure coefficient convey different information about the underlying relationship. The sign of the structure coefficient is interpreted with reference to the position of group centroids on the discriminant function.

discriminating power. Putting these two hypotheses together, we speculated that the discriminating power of knowledge items (as indexed by values of the structure coefficients) might be negatively correlated with the rated "knownness" of items. As described below, this speculation turned out to be only half correct.

We computed the correlation between question means on the "known" scale and the Fisher-transformed values of the structure coefficients.[27] The obtained values of Pearson's r were .47 (p=.0001), .58 (p=.0001), and .45 (p=.002) for battalion, company, and platoon-level data, respectively. Thus, at all three levels, high positive values of the transformed structure coefficients were associated with high "knownness" ratings and high negative values of the transformed structure coefficients were associated with low "knownness" ratings. To interpret these results we must reexamine the results of our discriminant analyses.[28]

Structure coefficients that were high and positive represented, in our study, items for which high goodness ratings were associated with high (novice-like) scores on the CDF. Thus, items with high positive structure coefficient values were items that were discriminating because experienced leaders liked them less than did novices. More precisely, items with high positive structure coefficients were items on which high goodness ratings were associated with an overall pattern of responding that was novice-like (i.e., close to the novice group centroid on the axis representing the CDF) and on which low goodness ratings were associated with an overall pattern of responding that was experienced-like.

Conversely, structure coefficients that were high and negative represented items for which high goodness ratings were associated with low (experienced-like) scores on the CDF. Thus, items with high negative structure coefficient values were items that were discriminating because experienced leaders liked them more than did novices. More precisely, items with high negative structure coefficients were items on which high goodness ratings were associated with an overall pattern of responding that was experienced-like (i.e., close to the experienced group centroid on the axis representing the CDF) and on which low goodness ratings were associated with an overall pattern of responding that was novice-like. Of course, items with structure coefficients in the intermediate range (of either sign) were nondiscriminating.

Having clarified the meaning of structure coefficient values in this study, we see that a significant positive correlation between transformed values of the structure coefficient and mean "knownness" ratings indicates two classes of highly discriminating tacit-knowledge items in our sample. What we will call Class A discriminating items are those that experienced leaders rated more highly than did

---

[27] Fisher's r-to-z transform yields a transformed value that is the inverse of the arc tangent of the original variable value. The transformation is commonly applied to distributions of correlation coefficients for purposes of hypothesis testing.

[28] For ease of exposition, the position of group centroids on the axis representing the CDF have been given a common interpretation across levels.

novices on the good scale. These items tended to be among the most well known knowledge items in the sample (as rated by all subjects). What we will call Class B discriminating items are those that novice leaders rated more highly than experienced leaders on the "good" scale. These items tended to be among the least well known knowledge items in the sample (as rated by all subjects). We speculate that Class A discriminating items may communicated and discussed but not fully appreciated until one has actually been in the job (i.e., led a platoon, company, or battalion). Class B discriminating items may reflect leader knowledge that is more particular in nature but only "apparently" good--knowledge that does not hold up well in actual practice and that experienced leaders tend to "see through" and reject. Because examples of both good and bad leader knowledge are needed to construct effective tacit-knowledge tests, discriminating items in both of the classes discussed above are likely to be useful in later test development.

Finally, note that we do not regard the relatively high "knownness" ratings of Class A discriminating items as evidence of their nontacitness. First, the degree to which knowledge is commonly known is only weakly related to its tacitness under the explanatory model of tacit knowledge. Second, half of our subjects (the novices at each level) may have been in a poor position to judge the degree to which knowledge is actually shared among leaders at a level above theirs in the chain-of-command. Finally, although mean "knownness" ratings for Class A discriminating items were among the highest in the sample, none of these means fell above 6.0 on the "known" scale (where a rating of 7 corresponded to the judgment "known by almost all"). Thus, "known" ratings for Class A discriminating items were high in relative rather than absolute terms.

In summary, our analysis of the TRADOC data yielded several important findings. First, we found that experienced and novice leaders, at each of the levels under study, showed patterns of goodness ratings to tacit-knowledge items that differed from one another significantly. Although these differences cannot be attributed solely to the effects of leadership experience on the basis of our data, they do suggest that the knowledge items in the TKS hold promise for development into tacit-knowledge tests that are comparably discriminating. Second, by computing an index of the discriminative power of items, we created a criterion on the basis of which the most promising items in the tacit-knowledge corpus can be selected for use in test construction. Table 6 shows the most promising items from the TRADOC analysis, organized by level, as measured by structure coefficient values from the discriminant analysis.

Table 6.
Top Decile of Tacit-Knowledge Items in Absolute Value of Structure Coefficient, by Level.

Level/ Question	Item
Battalion	
Q54	If you want to ensure that your battalion does not become overwhelmed by competing demands, then select three-to-five upcoming missions on which to focus the unit's attention. Focusing on more that five missions at once is unrealistic and counter-productive. (.30)
Q19	If you want to effectively develop your company commanders, then assign majors from the battalion staff to mentor them. Ensure that the mentor is from the same branch as the company commander. The non-rating relationship between majors and captains can facilitate open communication and promote development. (.30)
Q29	If your commander has unjustly destroyed the careers of several officers, unit morale seems dangerously low, and all attempts to communicate with the commander have failed, then by-pass the chain of command and speak with your boss's commander directly because such action is in the best interests of your unit. (.27)
Q13	If you are feeling lonely in your role as battalion commander, then find someone sympathetic to use as a "sounding board" for your ideas. Having a confidant provides you with feedback on your ideas and boosts your confidence. (-.25)
Q41	If you have the authority to make changes to the inside and outside areas of the billets, then take steps to make them more livable for single soldiers (e.g., planting a garden). You can increase the satisfaction of soldiers in the billets by caring for them in this way. (-.22)
Q9	If your boss makes a mistake in a public forum and does not ask for comments or questions, and if you feel a need to confront the boss about the mistake in public, then speak directly to the issue, avoiding evaluative statements that may embarrass your boss. Saving your boss from embarrassment helps to preserve your relationship with him or her. (-.21)

Table 6. Continued

Q32	If constant negative feedback about the performance of your unit causes you to become frustrated and if you have a good relationship with your CSM, then discuss your frustrations and feelings with him or her. Talking things through with the CSM may prevent you from venting your frustrations on your soldiers. (-.21)
Company	
Q12	If your staff NCOICs are not supporting company policies and standards, then formalize those policies and standards in the unit SOP. Once responsibilities are formally established in an SOP, the power of UCMJ action can be used to encourage compliance. (.44)
Q59	If you want your NCOs to exercise initiative, then give them mission-type orders and involve them in decision-making. These techniques encourage them to exercise initiative. (-.41)
Q58	If a sensitive item is lost and you are confident that the item will be found before the next reporting period and if there is a high level of trust within the unit, then hold-off on reporting the loss until the next reporting period while following SOPs to correct the situation and preparing for a timely report if the item is not found. Holding-off in this way gives you time to correct the problem and protect your unit from unnecessary bad publicity. (.37)
Q9	If you command a unit made up of both military and civilian personnel, and have encountered problems relating to the allocation of work between the two groups, then use a sign-out sheet to make visible each member's location during the work day. By making each employee's location visible to all others you prevent perceptions of unfair allocation of work. (.33)
Q21	If one of your male soldiers refuses to meet a training standard that he has the ability to meet, and you are a female commander, then challenge the soldier's male pride in order to get him to meet the standard. A female commander is in a unique position to use a male soldier's pride to motivate him. (.32)

Table 6. Continued

Q40	If your unit is dispersed among different garrison commanders who have non-judicial authority over your soldiers and control resources that improve your soldiers quality of life, then build a good relationship with those commanders by visiting them regularly. A good relationship with garrison commanders enables you to exercise indirect influence over your soldiers' quality of life. (-.32)
Q56	If your in-box has paperwork in it at the end of the day, then go through it and act on all time-sensitive matters related to soldier welfare (e.g., pay inquiries). Prompt action on soldier-related matters shows your soldiers that you care for them. (.31)
Platoon	
Q34	If your commander does not provide you with adequate performance feedback, then use your fellow lieutenants as a support group in order to get the performance feedback you need. (.48)
Q10	If normally good soldiers become negative, begin to ignore orders, or begin to talk back to their leaders, then try not to assign them another mission until they have had time to rest. These indicators suggest that your soldiers have reached their limits and must be rested in order to maintain their combat effectiveness. (.35)
Q44	If you need to correct a soldier, and the problem is not one of immediate urgency (e.g., a soldier who does not salute or honor retreat) or one requiring that an officer get involved, then communicate the correction through the NCO chain of command in order to preserve the differentiation between officer and NCO roles. (.34)
Q24	If your commander issues a directive with which you do not agree, and you are unable to persuade him or her to change the directive, then let your NCOs know how you really feel about the directive and then try to convince them to support it nonetheless. Talking straight with your NCOs maintains your credibility with them. (.33)

Table 6. Continued

Q2	If your commander issues a directive with which you do not agree and you feel that you understand his or her rationale, then use METT-T to frame an argument in favor of changing the directive. Showing how a directive will impact on the unit's mission increases the persuasiveness of your argument. (-.30)

Relationship of Tacit Knowledge to Effectiveness (FORSCOM)

Summary statistics.

Sample sizes for the FORSCOM data collection are shown in Table 7. Sample sizes for TKS and LES administrations are shown separately. Values in this table, for both TKS and LES samples, represent the number of subjects (i.e., individuals about whom information was sought) rather than the number of respondents (i.e., individuals providing information about subjects). As stated earlier, respondents and subjects are the same in the TKS data. Note that decreasing sample sizes across levels (from platoon to company to battalion) simply reflect the decreasing number of incumbents at successive levels within the chain-of-command.

Table 7.
Sample Sizes for FORSCOM Sample, by Survey Form and Level

Survey	Level		
	Battalion	Company	Platoon
TKS	28	132	287
LES			
Superior	34	133	296
Peer	---	174	311
Self	20	128	195
Subordinate	37	109	---

Summary statistics for ratings on the four TKS scales are shown in Table 8. Mean values in this table are grand means computed across questions and subjects. Standard deviations reflect the dispersion among question means. Thus, mean values in this table are based upon 66, 67, and 46 question means, for battalion, company, and platoon-level data respectively. For all of the levels under study, the highest

standard deviations were observed for the good and concept scales, with the exception of the battalion level for which the highest standard deviation was observed for the often scale. Between-level differences on the good scale were nonsignificant in a Kruskal-Wallis nonparamteric test on ranks ($X^2(2)=3.34$, p=.18).

Table 8.
Means and Standard Deviations (in parentheses) on TKS Ratings, by Level (FORSCOM Sample).

Rating Scale	Level		
	Battalion	Company	Platoon
Good	5.51 (.69)	5.29 (1.17)	5.19 (.88)
Known	5.02 (.52)	4.87 (.62)	4.53 (.52)
Often	4.42 (.88)	4.70 (.90)	4.59 (.65)
Concept	5.18 (.79)	5.10 (1.17)	4.93 (.95)

Summary statistics for ratings on the four LES rating scales are shown in Table 9. Mean values in the table are computed across subjects. Standard deviation values reflect the dispersion among subject means.

Table 9.
Means and Standard Deviations (in parentheses) on LES Ratings, by Level (FORSCOM Sample).

Rating Scale	Level		
	Battalion	Company	Platoon
Superior	2.65 (.92) n=34	2.56 (1.08) n=133	2.97 (1.00) n=296
Peer	---	3.08 (.96) n=174	3.12 (.97) n=311
Self	2.75 (.85) n=20	2.76 (.79) n=128	2.84 (.87) n=195
Subordinate	2.93 (.97) n=37	2.92 (.96) n=109	---

Rating-scale intercorrelations.

The intercorrelations among rating scales in the TKS are shown in Table 10. Correlation coefficients were computed on question means in order to remove the effects of within-subject response dependencies. All correlation coefficients were highly significant ($p \leq .0001$ in all cases).

Table 10.
Intercorrelations Among TKS Rating Scales, by Level (FORSCOM Sample).

	Mean Good	Mean Known	Mean Often	Mean Concept
Mean Good				
Battalion				
Company				
Platoon				
Mean Known				
Battalion	.86			
Company	.89			
Platoon	.89			
Mean Often				
Battalion	.50	.51		
Company	.66	.69		
Platoon	.61	.67		
Mean Concept				
Battalion	.94	.86	.61	
Company	.99	.92	.67	
Platoon	.99	.93	.64	

$p \leq .0001$ in all cases.

As with the TRADOC data, a principal components analysis was performed on the correlation matrix at each level. Again, each such analysis yielded second and subsequent principal components that accounted for less than one variable's worth of variance, indicating a single, general, "quality" factor. As in the TRADOC analyses, and following the same logic as that described in an earlier section, we decided to focus on goodness ratings in later analyses.

The interrelation among effectiveness scales in the LES is shown in Table 11. Correlation coefficients were computed on subject means. That is, each r-value is based on n sets of paired mean ratings applied to a given subject (e.g., the mean peer rating applied to subject i and the mean subordinate rating applied to subject i). Of course, for correlations involving the superior and self scales, "mean" subject ratings were each based on a single observation. As the table shows, correlation coefficients tended to be significant with the exception of several correlations involving the self scale (i.e., subject's ratings of their own leadership effectiveness). Subject means on the self scale were correlated at non significant levels with superior and subordinate ratings (for battalion commanders) and with superior ratings (for platoon leaders).

Table 11.
Intercorrelations Among LES Rating Scales, by Level (FORSCOM Sample).

	Superior	Self	Peer	Subordinate
Superior				
Battalion				
Company				
Platoon				
Self				
Battalion	.03			
Company	.25*			
Platoon	.10			
Peer				
Battalion	---	---		
Company	.49*	.26*		
Platoon	.25*	.21*		
Subordinate				
Battalion	.55*	-.02	---	
Company	.38*	.24*	.39*	
Platoon	---	---	---	

* $p \leq .02$

Correlations between item ratings and effectiveness.

Leader effectiveness scales were examined (separately) in terms of their correlation with goodness ratings on tacit-knowledge items. The purpose of these analyses was to identify those tacit-knowledge items for which goodness ratings may be related to the rated effectiveness of leaders. For battalion commander data, scores on each of three effectiveness scales (superior, self, subordinate) were correlated with goodness ratings on each of 66 items, for a total of 198 correlations. For company commander data, scores on all four effectiveness scales (superior, peer, self, subordinate) were correlated with goodness ratings on each of 67 items, for a total of 268 correlations. For platoon leader data, scores on each of three effectiveness scales (superior, peer, self) were correlated with goodness ratings on each of 46 items, for a total of 138 correlations. The percentage of correlation coefficients that met conventional standards of significance ($p \leq .05$) were as follows: 6% for battalion, 7% for company, and 15% for platoon.[29] Thus, we obtained evidence of an association between item ratings and effectiveness, for some items, at

[29] Significance tests based on the quantity $r(n-2)^{1/2}/(1-r^2)^{1/2}$ distributed as *t* for two degrees of freedom.

each of the levels under study. Table 12 shows the distribution of significant correlations across the fours effectiveness scales.

Table 12.
Number and Proportion (in parentheses) of Significant Correlations, by Effectiveness Scale and Level.

Rating Scale	Level		
	Battalion	Company	Platoon
Superior	1 (.07)	3 (.16)	2 (.09)
Peer	---	3 (.16)	7 (.33)
Self	7 (.53)	8 (.42)	12 (57)
Subordinate	5 (.38)	5 (.26)	---

The large number of statistical tests reported above raises concern regarding an inflated probability of family-wise Type I error. That is, although the per-comparison probability of Type I error is fixed at .05, the probability of at least one Type I error among the entire set of comparisons is essentially unity according to standard methods of estimation (Hays, 1988). Using corrective methods such as those based on the Bonferroni inequalities we may restrict the probability of family-wise Type I error to some lower value but, given the very large number of correlations, the resulting decision rules are extremely stringent (e.g., reject H_0 only if $p \leq .0003$) and we essentially trade Type I error for Type II error.

When we consider the relative cost of the two types of errors within the context of the current study, we see that it would be unwise to inflate the probability of Type II error in order to ward off the acceptance of spurious correlations. Items correlated with effectiveness at the population level, if excluded from further consideration due to Type II error, are lost to the test-development process. Items uncorrelated at the population level, if not excluded from further consideration due to Type I error, will have ample opportunity to be rejected during the course of item construction and test validation. In light of the relatively low cost of Type I error, we will regard items significant at the .05 level (uncorrected) as promising for purposes of instrument development.

Correlations between item ratings and effectiveness (point-biserial).

Leader effectiveness was again examined in terms of its correlation with goodness ratings on tacit-knowledge items. In these analyses, subjects were divided into quartiles on each of the effectiveness scales. High and low effectiveness groups were formed from the top and bottom quartiles on each effectiveness measure. A

point-biserial correlation coefficient was computed between goodness ratings on tacit knowledge items in the TKS and a categorical variable representing effectiveness group (high versus low). The purpose of these analyses was to identify those tacit-knowledge items for which goodness ratings may be related to the rated effectiveness of leaders, but to do so under conditions of the clearest contrast in leader effectiveness.

The percentage of correlation coefficients that met conventional standards of significance ($p \leq .05$) were as follows: 8% for battalion, 9% for company, and 16% for platoon. Thus, we again observed an association between item ratings and effectiveness, for some items, at each of the levels under study. Table 13 shows the distribution of significant correlations across the four effectiveness scales.

Table 13.
Number and Proportion (in parentheses) of Significant Correlations, by Effectiveness Scale and Level (Point-Biserial).

Rating Scale	Level		
	Battalion	Company	Platoon
Superior	6 (.37)	3 (.12)	2 (.09)
Peer	___	5 (.21)	8 (.36)
Self	5 (.31)	8 (.33)	12 (.54)
Subordinate	5 (.31)	8 (.33)	___

In summary, our analysis of the FORSCOM data yielded several important findings. First, we found significant relationships between item ratings and leader effectiveness for a number items at each level under study. Although these relationships cannot be attributed solely to individual differences in effectiveness on the basis of our data, they do suggest that some of the knowledge items in the TKS hold promise for development into tacit-knowledge test questions that discriminate between more effective and less effective leaders. Tables 14 through 17 show the most promising items from the FORSCOM analysis, organized by level. Promisingness is here indexed by the absolute value of the Pearson product moment correlation between goodness ratings on items and effectiveness ratings made by superiors, peers, self, and subordinates. Items from the top decile on this index are shown in the tables.

Table 14.
Top Decile of Tacit-Knowledge Items in Absolute Value of Correlation Coefficient (Good X Superior), by Level.

Level/ Question	Item
Battalion	
Q13	If you are feeling lonely in your role as battalion commander, then find someone sympathetic to use as a "sounding board" for your ideas. Having a confidant provides you with feedback on your ideas and boosts your confidence. (.42)
Q42	If you want to develop your subordinate commanders, then talk to their soldiers in order to determine the areas in which the commanders are weak. Because soldiers directly experience the consequences of a commander's weaknesses, they are a good source of information. (.39)
Q11	If you want to correct a perception among your soldiers that you do not support them on award approvals, then invite a few soldiers to observe your decision-making process and encourage them to share what they learned with their peers. By showing soldiers some of the complexities of the approval process, you can correct misperceptions regarding fairness. (.37)
Q50	If you want to develop company-grade officers, then focus on developing their thinking skills through brief-backs and by monitoring their participation in a professional reading program. (-.37)
Q38	If you want to build trust with your soldiers, then trust your soldiers to do their jobs, consistently maintain your composure when problems arise, be accessible to your soldiers, help them solve problems, and take risks to protect their welfare (e.g., questioning your commander about unreasonable taskings). The above behaviors help a leader build trust with his/her subordinates. (-.36)

Table 14 Continued

Q48	If senior NCOs with needed expertise are undermining the authority of a company commander, then develop the expertise of junior NCOs while getting rid of the subversive NCOs through attrition. This course of action eliminates the expert power base of the subversive NCOs and helps you maintain control of your unit. (.33)
Q62	If you have bypassed the chain-of-command in order to resolve a problem with your commander and if this event has created dissension within your own unit, then conduct open sensing sessions with your subordinate leaders in order to correct any misperceptions as well as to demonstrate your own availability and openness. (.32)
Company	
Q36	If you decide to modify the TO&E chain of command to place an officer in charge of a section, then place the officer in the rating chain of the section NCOIC. This technique allows the officer to hold the NCOIC accountable. (.27)
Q66	If a training event scheduled by your battalion commander (e.g., ADA Battalion) conflicts with a training event scheduled by your supported-unit commander (e.g., Infantry Brigade Commander) and if the two events have equal training value and impact on soldiers' quality of life, then support the training event scheduled by your battalion commander. By doing so you show loyalty to your battalion commander. (.21)
Q48	If you want to encourage soldiers to take initiative, then give them responsibility for the daily affairs of the unit and praise/reward them on the basis of initiative. Rewarding soldiers for taking initiative increases the likelihood that they will do so in the future. (.20)
Q57	If you want to persuade soldiers to take precautions during off-duty hours (e.g., against DUI, drowning, mugging), then ask a soldier who has been a victim of one of these hazards to speak to the unit about his or her experience. By using a soldier who has had actual experience of a hazard, you increase the persuasiveness of the message. (.19)

Table 14 Continued

Q15	If you are deployed on an FTX, then inspect the technical and tactical aspects of your soldiers' performance that relate to their individual survival (e.g., range cards and fighting positions). Interest in these matters shows that you care for your soldiers' welfare. (-.18)
Q62	If you are deployed and sustaining continuous operations, then take time out each day to read or contemplate in order to reduce stress and preserve your mental effectiveness. (.17)
Q8	If you have a major decision to make, then let your junior officers in on the decision-making process in order to give them a feeling of ownership of policies and train their thought processes. (.16)
Platoon	
Q2	If your commander issues a directive with which you do not agree and you feel that you understand his or her rationale, then use METT-T to frame an argument in favor of changing the directive. Showing how a directive will impact on the unit's mission increases the persuasiveness of your argument. (-.18)
Q46	If you have an insubordinate NCO whom your commander refuses to discipline for what you feel are racial reasons, then confront the commander and, if that fails, take the matter up with the battalion XO. Going to the XO minimizes the impact of going over your commander's head. (.15)
Q3	If you are taking charge of a new unit and are unsure of yourself, then be careful to present a confident image while you get yourself up to speed. Don't pretend to know what you don't know and state what you do know with conviction. In this way you inspire trust in others and open the flow of communications. (-.13)
Q42	If your unit fails because of a mistake you made, and if that mistake is made public in an AAR, then reflect on the mistake to determine the lesson learned but put it behind you as soon as possible. By learning from rather than dwelling on mistakes you improve as a leader. (-.13)

Table 14 Continued

Q13	If your soldiers are deployed in a hostile theater but not actively engaged with the enemy, continue to require garrison-like routines (e.g., gear stowed beneath bunks, boots shined, etc.). This practice helps soldiers to retain some feeling of normalcy in their lives and maintains discipline in the unit. (-.12)

Table 15.
Top Decile of Tacit-Knowledge Items in Absolute Value of Correlation Coefficient (Good X Peer), by Level.

Level/ Question	Item
Company	
Q18	If you want to encourage initiative and risk taking among your soldiers, but you feel that they have been punished for it in the past, then model initiative and risk taking in promoting the welfare of the unit (e.g., painting unit logos on vehicles without permission). Your modeling of initiative and risk taking will build trust with soldiers and encourage them to exercise initiative and take risks. (-.24)
Q37	If a subordinate leader fails in an important mission but you are not physically present and do not know the details, then gather information before taking action. After determining the cause of the failure, brief your commander on the problem and how to prevent it in the future. By gathering information before acting, you preserve your relationship with the junior leader. (.20)
Q43	If you have established a climate of trust with your soldiers, then you can withhold information from them, on occasion, without negative consequences. Soldiers are willing to accept the withholding of information if they trust the commander's motives. (-.20)
Q31	If you want to communicate to your soldiers that you are displeased with them but do not want to lose your composure then use nonverbal methods such as facial expressions or becoming more formal in your demeanor. These nonverbal methods can communicate displeasure more effectively than ranting and raving at soldiers. (-.17)
Q29	If you need feedback from your soldiers about the unit, then take advantage of informal settings to talk to them (e.g., talk to them while eating in the back of a track) or create an informal setting in your office by arranging chairs in a circle. When soldiers feel relaxed, you receive more candid feedback from them. (.16)

Table 15 Continued

Q62	If you are deployed and sustaining continuous operations, then take time out each day to read or contemplate in order to reduce stress and preserve your mental effectiveness. (.17)
Q44	If you anticipate difficulties in counseling a subordinate, then role play the counseling beforehand and prepare responses to possible counselee reactions. Rehearsal and anticipation of objections increases your confidence and ability to control the situation. (.15)
Platoon	
Q24	If your commander issues a directive with which you do not agree, and you are unable to persuade him or her to change the directive, then let your NCOs know how you really feel about the directive and then try to convince them to support it nonetheless. Talking straight with your NCOs maintains your credibility with them. (.22)
Q45	If you provide support to another unit but the personnel in that unit do not fulfill their role obligations, then take on extra responsibility in order to make it easy for the supported unit to accept your support (e.g., pick up broken radios instead of waiting for the supported unit to bring them to you). The easier you make it to accept your support, the greater your ability to add value to the organization. (.22)
Q4	If you are worried that a suicidal soldier in your unit may be teased or isolated by other soldiers, and if this concern outweighs the soldier's right to privacy, then hold a meeting in which you ask the members of the unit to help you care for the troubled soldier by watching over him and by refraining from teasing him. This approach prevents the soldier's problems from being compounded by others in the unit. (.22)
Q38	If you are taking over a unit with combat experience but lack such experience yourself, then listen to your soldiers, speak to them in a tone that conveys respect, and do not change procedures that work. These activities will allow you to build a relationship with combat veterans. (.17)

Table 15. Continued

Q28	If your unit is deployed for an extended period in a remote area, then keep your soldiers busy with PT and cross-training with other branches. These measures keep soldiers from becoming bored. (.16)

Table 16.
Top Decile of Tacit-Knowledge Items in Absolute Value of Correlation Coefficient (Good X Self), by Level.

Level/ Question	Item
Battalion	
Q4	If you experience anger at poor unit performance, then seek additional information from experienced soldiers in the unit before acting on your anger. This technique allows you to consider situational factors when arriving at an account of the failure. (-.74)
Q45	If you wish to give a deserving officer a command but have doubts about the officer's chances for success, then put the officer in command of a unit with strong subordinate leaders as this will protect the organization and maximize the probability of a successful command. (.63)
Q16	If you plan to give company-level command to an executive officer, then place him or her in command of a unit in which he or she was not "raised." By putting the officer in command of a new unit, you force him to assess the strengths and weaknesses of that unit. (-.63)
Q64	If a company commander is not supporting your vision, then initially have the battalion XO inform him or her that there is a problem. Using an indirect approach in this situation allows you to preserve a positive relationship with the errant company commander. (-.61)
Q3	If a commander's unit is not performing, if soldiers and their families lack confidence in the commander, and if the commander is incompetent and fails to show improvement, then relieve the commander. Such an officer is not fit to command and must be relieved in order to protect the organization. (-.59)
Q60	If your company commanders are ambitious and have a tendency to take on resource-intensive missions that exceed the capacities of their units, then require them to conduct a resource assessment before they take on missions. By requiring a resource assessment you prevent commanders from taking on missions that may over burden their units. (-.57)

Table 16 Continued

Q59	If a subordinate requests a transfer due to family problems but you feel that work conditions do not contribute to the family problems, then deny the transfer. You take care of soldiers when you make them face up to their problems. (-.56)
Company	
Q53	If you give a soldier an operation order and want to be sure that he or she understands the order, then pay attention to the sequence in which the soldier recalls events during a brief-back. Recall sequence provides you with an indication of how well the information is understood. (-.27)
Q26	If your unit is dispersed and assigned to different garrison commands and weekly training meetings require subordinate leaders to travel a substantial distance, then ask a different platoon leader to host the training meeting each month. This technique saves your subordinates driving time and provides an opportunity for platoon leaders and PSGs to share their ideas on how to run their units. (-.26)
Q17	If you are required to cross-attach one of your platoons during a major training exercise, and if you feel that one of your platoons is not fully trained, then cross-attach your best platoon. You are better prepared to deal with a weak platoon than is a commander from another branch. Cross-attaching your best platoon to outside units builds and maintains trust in your organization. (-.24)
Q15	If you are deployed on an FTX, then inspect the technical and tactical aspects of your soldiers' performance that relate to their individual survival (e.g., range cards and fighting positions). Interest in these matters shows that you care for your soldiers' welfare. (-.23)
Q14	If a junior officer is insubordinate to you in private, then immediately reprimand the officer in order to protect your authority. (-.23)

Table 16 Continued

Q38	If you want to find out whether or not your message (e.g., regarding safety precautions) is getting through to your soldiers, then select a soldier at random and ask for a spontaneous brief-back on the topic. A spontaneous brief-back provides a good means to check understanding and retention of your message, as well as incentive for soldiers to listen closely. (-.22)
Q58	If a sensitive item is lost and you are confident that the item will be found before the next reporting period and if there is a high level of trust within the unit, then hold-off on reporting the loss until the next reporting period while following SOPs to correct the situation and preparing for a timely report if the item is not found. Holding-off in this way gives you time to correct the problem and protect your unit from unnecessary bad publicity. (-.20)
Platoon	
Q27	If soldiers question your authority or show evidence of misplaced priorities in their approach to a mission, then use directive leadership to correct them. Indirect methods are not appropriate under these circumstances. (-.26)
Q22	If you are trying to change your own behavior in order to become a more effective leader, then put visible reminders in places where you will see them throughout the day. These notes will help you to monitor behaviors that you are seeking to change. (-.22)
Q37	If your commander issues a directive with which you do not agree and you are unable to persuade him or her to change the directive, then discuss the problem with other key leaders in the unit and arrange to meet with the leader to discuss the problem as a group. A large group of subordinate leaders, all holding the same position, may persuade the commander to reconsider. (.20)
Q11	If frequent changes to the battalion training schedule have an adverse effect on your soldiers' morale, then publish your own training schedule, based on the battalion schedule, that is short term but specific. A platoon-level training schedule will make your soldiers feel informed rather than jerked around. (-.19)

Table 16 Continued

Q41	If your commander makes a decision with which you do not agree, then frame your input as a request for guidance and communicate the impact of the decision on the unit. Adopting a less confrontational approach can help maintain open communications with the commander. (-.18)

Table 17.
Top Decile of Tacit-Knowledge Items in Absolute Value of Correlation Coefficient (Good X Subordinate), by Level.

Level/ Question	Item
Battalion	
Q41	If you have the authority to make changes to the inside and outside areas of the billets, then take steps to make them more livable for single soldiers (e.g., planting a garden). You can increase the satisfaction of soldiers in the billets by caring for them in this way. (.49)
Q48	If senior NCOs with needed expertise are undermining the authority of a company commander, then develop the expertise of junior NCOs while getting rid of the subversive NCOs through attrition. This course of action eliminates the expert power base of the subversive NCOs and helps you maintain control of your unit. (.49)
Q42	If you want to develop your subordinate commanders, then talk to their soldiers in order to determine the areas in which the commanders are weak. Because soldiers directly experience the consequences of a commander's weaknesses, they are a good source of information. (.47)
Q63	If you take responsibility for a task away from a leader and give it to one of his or her subordinates, replace the task with one of equal or greater responsibility. Replacing the task communicates trust to the leader and sustains his/her feeling of competence. (.40)
Q19	If you want to effectively develop your company commanders, then assign majors from the battalion staff to mentor them. Ensure that the mentor is from the same branch as the company commander. The non-rating relationship between majors and captains can facilitate open communication and promote development. (.39)

Table 17 Continued

Q49	If one of your subordinates makes a mistake in a public setting, do not humiliate him or her. Rather, use the mistake as an opportunity to develop the officer in private, emphasizing recognition of the problem and corrective action. Discussing mistakes in a nonthreatening way promotes learning and development. (.39)
Q66	If you have taken steps to make your soldiers' living areas more pleasant and if soldiers from other units express an interest in your efforts, then extend access to these soldiers. Taking care of soldiers does not stop at the boundaries of your unit. (.39)
Company	
Q18	If you want to encourage initiative and risk taking among your soldiers, but you feel that they have been punished for it in the past, then model initiative and risk taking in promoting the welfare of the unit (e.g., painting unit logos on vehicles without permission). Your modeling of initiative and risk taking will build trust with soldiers and encourage them to exercise initiative and take risks. (-.35)
Q60	If you are a female commander and your male NCOs have trouble taking orders from you, then involve them in decision-making. This technique gets NCOs to commit to a course of action and avoids resistance to authority based on gender. (-.28)
Q7	If any of your soldiers are having marital problems, then offer space in the billets to married soldiers. This proactive measure helps soldiers to avoid trouble (i.e., domestic disturbance) and protects their families. (-.28)
Q16	If you want to identify the informal leaders in each squad, then determine who the soldiers themselves seek out for advice. Soldiers will usually seek advice from the few people who have actually read the TMs and FMs. These informal leaders are often good sources of information about problems in the unit. (-.28)
Q39	If you want to use a mission as an opportunity to develop junior officers, then explain the "big picture" of the mission to them rather than simply their individual piece of the mission. (-.24)

Table 17. Continued

Q26	If your unit is dispersed and assigned to different garrison commands and weekly training meetings require subordinate leaders to travel a substantial distance, then ask a different platoon leader to host the training meeting each month. This technique saves your subordinates driving time and provides an opportunity for platoon leaders and PSGs to share their ideas on how to run their units. (-.19)
Q40	If your unit is dispersed among different garrison commanders who have non-judicial authority over your soldiers and control resources that improve your soldiers quality of life, then build a good relationship with those commanders by visiting them regularly. A good relationship with garrison commanders enables you to exercise indirect influence over your soldiers' quality of life. (-.19)

Finally, note that responses to the relatively decontextualized knowledge items in the TKS represent a weak test of the ultimate predictive validity of tacit-knowledge test questions. According to the explanatory model of tacit knowledge, the more "event-like" the performance measure the greater should be the salutary effects of having acquired relevant tacit knowledge. Because the knowledge items in the TKS were not very event-like, we did not expect these items to discriminate, on the basis of tacit knowledge possession, to the same degree as would more fully contextualized tacit-knowledge test questions which will include longer scenario descriptions and a greater range of response alternatives. For this reason, the issue of predictive validity is addressed in a preliminary, rather than conclusive, way by the correlational analyses described above.

Organization of Tacit Knowledge (USMA)

Non-metric scaling.

The goal of the USMA data collection was to obtain preliminary evidence regarding the internal structure of the construct domain. A possible source of evidence regarding this structure is the organization of tacit knowledge in the minds of military leaders. Thus, we conducted a series of analyses in which proximity matrices from the USMA sort data were fit to a nonmetric multidimensional scaling model (MDS in what follows). Briefly, the MDS model takes as input a set of interobject proximities (in our case, the matrix of co-occurrence between tacit-knowledge items, collapsed across subjects) and derives a set of interobject distances--in Euclidean space of specified dimensionality--that best recovers the input proximities. The derived distances (and the set of j dimensions on which these distances are defined) are taken to represent the latent structure in the proximity

data. In the scaling of sort data such as ours, the derived distances and dimensions are taken to reflect the organization of knowledge in memory.

In MDS, the level of fit between derived distances and input proximities is called "stress" and lower stress values signal better levels of fit. The MDS model-fitting procedure is iterative, with successive iterations producing decrements in stress and with iteration terminating when stress stabilizes. The number of dimensions in a completed MDS solution is determined by examining both final stress levels and substantive interpretability for a range of dimensionalities. At each of the levels under study, we considered up to eight dimensions, using the stress-by-dimensionality plot, along with rough judgments of interpretability, to settle on five dimensional, five dimensional, and four dimensional solutions for battalion, company, and platoon-level data respectively. At these dimensionalities, we achieved very good levels of fit, given the number of objects being scaled (final stress values were .1034, .1244, and .1336 for battalion, company, and platoon-level solutions, respectively).

Interpretation of dimensions

At a given level, the key output of the model-fitting procedure was a list of coordinate values for each object or item on each dimension. Thus, the position of object i in j-dimensional space was represented by j real values corresponding to object i's position on dimensions 1 through j. In order to interpret the meaning of each dimension, we asked a panel of expert judges to apply a label to each of the derived dimensions at each of the levels under study. The task called for judges to examine the relative position of knowledge items on each of the dimensions and, while simultaneously considering the content of items, apply meaningful labels to the dimensions. The process of labeling dimensions derived in MDS model fitting is similar, in practice, to the labeling of factors derived in factor analysis. That is, both dimensions and factors are constructs defined by the position or "loading" of objects upon them. By asking judges to interpret the meaning of derived dimensions, we hoped to construct a model of the semantic relationships among leadership tacit knowledge in the minds of our sorters.

In order to facilitate the labeling of dimensions, multiple copies of the tacit-knowledge items were created and each copy was ordered according to the values of object coordinates on a given dimension. Thus, when judges assembled to interpret the derived dimensions, they were able to view the content, absolute position, and relative position of items on each dimension. The dimensions, represented in this fashion, were arranged around the tables and walls of a large classroom, with the dimensions for each level assigned to a separate area of the room. Each of three expert judges visited the room and studied the dimensions on two occasions--once individually and once in the presence of the other two judges. During the first session, each judge formed a personal impression concerning the meaning of each dimension. During the second session, judges met, reviewed the content and position of items on dimensions, discussed the merits of various dimension labels, and reached consensus on a single set of labels.

In assembling the panel of expert judges, we sought to include Army officers who were experienced at each of the levels under study. Consequently, we selected officers who had successfully commanded at all three levels under study and who had either completed or were currently working in colonel-level command positions. Our panel of judges consisted of two Colonels and a senior Lieutenant Colonel. Instructions to judges are shown in Appendix C.

The structure of the knowledge space.

The labels that the expert judges applied to each of the dimensions derived from MDS model fitting are shown in Tables 18, 19, and 20 (for battalion, company, and platoon levels, respectively). As the contents of these tables show, judges did not produce traditional, bipolar interpretations of the derived dimensions. Rather, their consensus labels reflected something akin to factor labels, with the positive and negative poles of each dimension corresponding, apparently, to degree of participation in the general theme established by high "loading" items (i.e., objects with high coordinate values on the dimension in question). Thus, according to our judges, the knowledge space derived from sort data is structured, at each level, by salient themes in leadership practice, with each knowledge item's position within the knowledge space reflecting its relevance or strength with respect to each of the underlying themes.

Table 18.
Dimensions of Battalion Commander Tacit Knowledge.

Dimension	Label	Explanation
1	Communicating a vision	Communicating goals by describing a future end state; including in that message issues of character, moral fortitude, and tough love (doing the right thing for the sake of the subordinate)
2	Establishing a climate for development	Communicating a set of beliefs or attitudes that allows subordinate development; reinforcing the statements by providing a structure of activities that supports such development
3	Managing the leader and the subordinate	Managing oneself while simultaneously "managing by exception" the problems that occur within the organization; considering the actions the leader should take to establish subordinate trust in the culture/climate/vision that has been communicated
4	Providing constancy	Providing stability by reinforcing the desired end state at every opportunity; communicating and maintain a uniform "commanders' intent"
5	Using influence tactics	Providing structure that allows subordinates to achieve desired levels of performance; maintaining authority by employing the full range of influence tactics; establishing parameters (in the form of formal controls) that reinforce subordinates' trust in core values

Table 19.
Dimensions of Company Commander Tacit Knowledge.

Dimension	Label	Explanation
1	Caring for soldiers through task completion	Knowing your job and making subordinate soldiers "do the right thing" (in terms of training readiness and task accomplishment) as a means of demonstrating to them that the leader cares for them
2	Prioritizing and solving problems	Dealing with day-to-day problems; communicating priorities and providing guidance to solve problems
3	Proactive decision-making	Thinking ahead to anticipate problems; sharing information so that subordinates can assist in proactive problem solving
4	Assessing risk	Determining the potential liabilities of an action; using team building to identify and potentially reduce hazardous situations
5	Short-term decision making	Providing face-to-face directions to influence an action at a critical moment; making decisions that facilitate day-to-day operations

Table 20.
Dimensions of Platoon Leader Tacit Knowledge.

Dimension	Label	Explanation
1	Acquiring confidence in interpersonal skills	Learning how to motivate subordinates; overcoming individual hesitancies towards motivating more experienced soldiers
2	Defining leadership style	Understanding one's personal leadership style; knowing the type of influence to use in one-on-one situations
3	Taking a stand	Confidently demonstrating concern for the units' welfare with subordinates; being forthright when discussing the strengths and weaknesses of the unit; acting for the benefit of the unit†
4	Taking and fostering accountability	Identifying problems (interpersonal or technical) within the unit and proactively seeking solutions to the problem; requiring the same actions of subordinates

† Judges felt that these actions may result in an attribution of selfishness.

To reiterate, the purpose of the USMA data collection, and subsequent scaling analyses, was to support later instrument development by exploring the internal structure of our knowledge sample. That is, we sought to go beyond the identification of promising items to the identification of dimensions that recover or reflect the similarity relations among all items. The dimensions outlined above, taken together, constitute our current best guess regarding the internal structure of the construct domain (the space of tacit knowledge for Army leadership) and, as such, may help us to construct tacit-knowledge tests whose composition reflects the range of construct-relevant knowledge.

The dimensions of leader knowledge outlined in Tables 18 through 20 represent an improvement upon classification schemes derived in earlier phases of the project (see Horvath et al., 1994a, 1994b). The dimensional representation, unlike the hierarchical-cluster representation upon which our previous classification scheme was based, specifies degree of relatedness between each item and each dimension. More importantly, the dimensional representation subsumes the

previous classification scheme which was based on a subset of the sort data from which the dimensional representation was derived.

Figures 4 through 6 illustrate the location of promising items in tacit-knowledge space. Figure 4 shows a sample two-dimensional projection from the five-dimensional knowledge space for battalion commanders. Data points are promising items of battalion commander knowledge (those in the top ten percent on at least two indices of promisingness) labeled by question number in the TKS. Figures 5 and 6 show similar projections from the company and platoon-level knowledge spaces, respectively.

Figure 4.
Location of Promising Tacit-Knowledge Items in a Sample Two-Dimensional Projection From the Battalion Commander Knowledge Space.

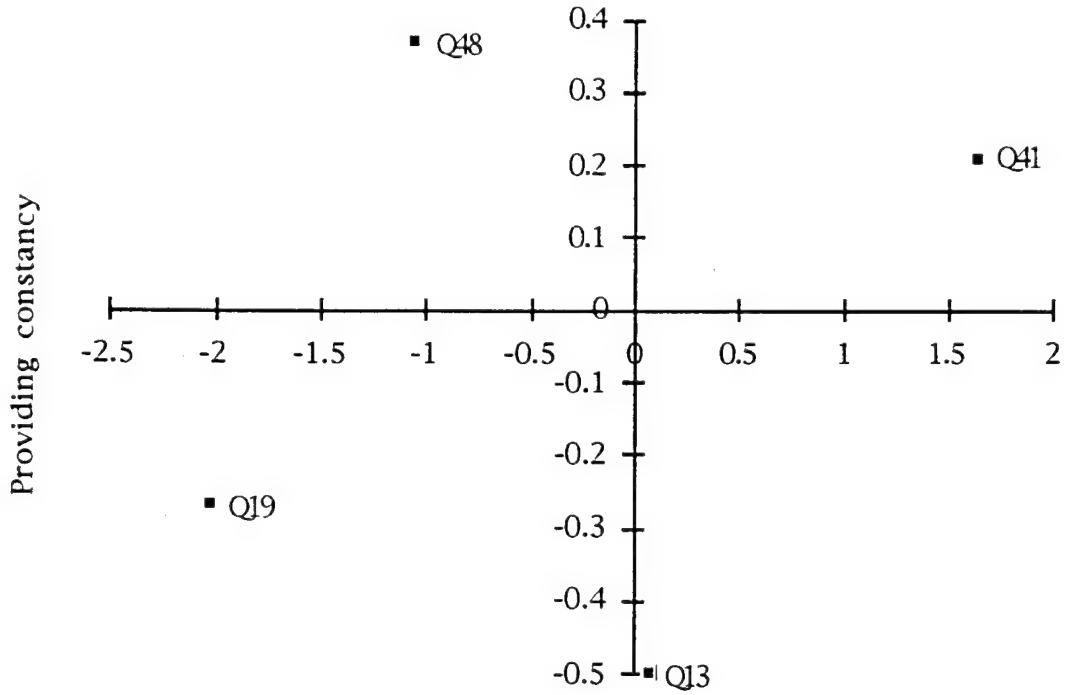


Figure 5.
Location of Promising Tacit-Knowledge Items in a Sample Two-Dimensional Projection From the Company Commander Knowledge Space.

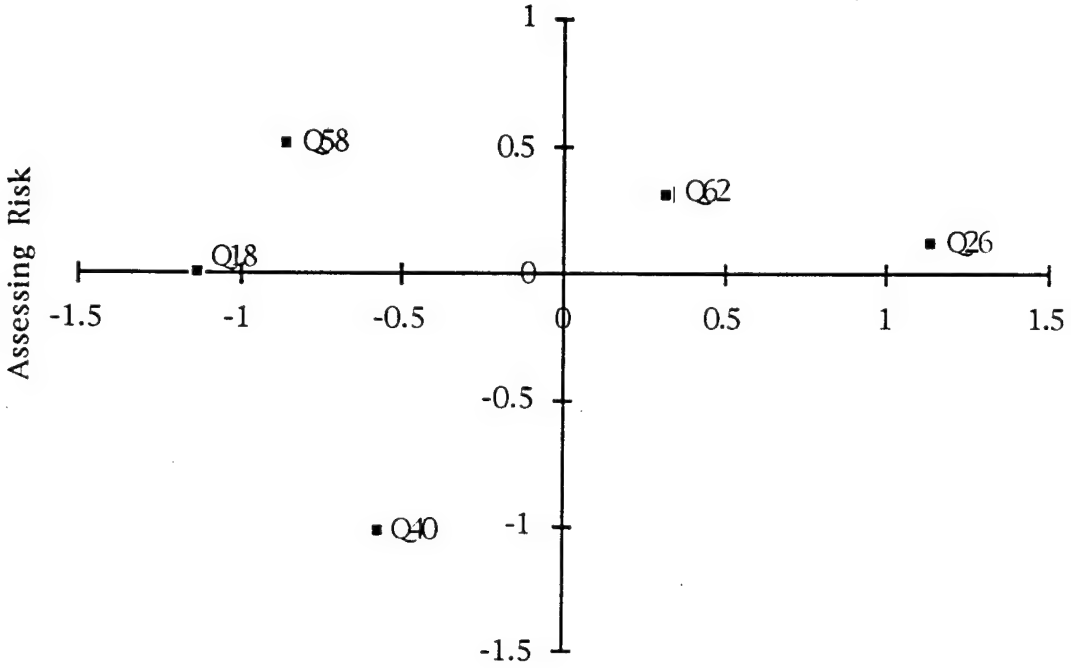
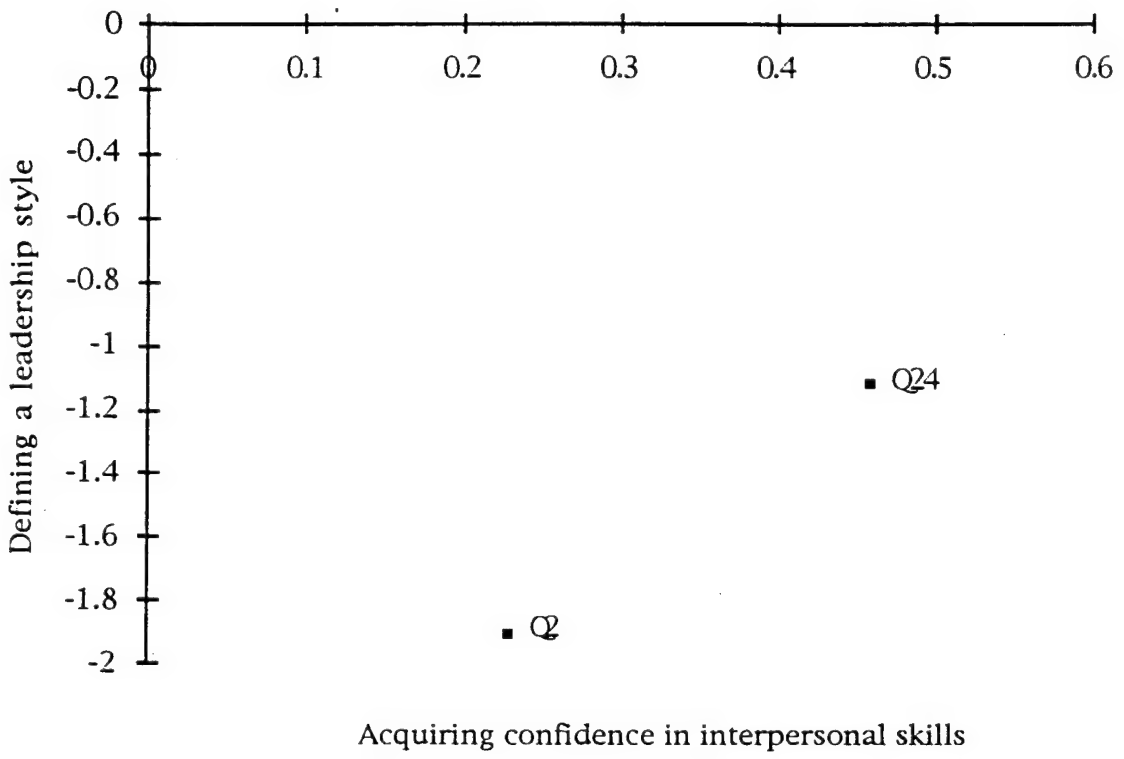


Figure 6.
Location of Promising Tacit-Knowledge Items in a Sample Two-Dimensional Projection From the Platoon Leader Knowledge Space.



As Figures 4 through 6 illustrate, the spatial representation of tacit knowledge will enable us to examine the distribution of promising items across semantically important dimensions of leadership knowledge during the process of test construction. This capacity to examine item distributions will provide an important supplement to "external" indices of promisingness in our efforts to construct tacit-knowledge tests that measure the focal construct.

Discussion

In the final section of this report, we seek to relate our theoretical and empirical work back to the goals of the study. We first revisit the tacit-knowledge construct in order to clarify the nature of tacit knowledge items and the nature of the tacit-knowledge tests based on those items. We then describe our general approach to the selection of items for test development. Finally, we seek to fit the results of the current study into an emerging validity profile for our to-be-constructed tacit-knowledge tests.

Tacit Knowledge as a Measurement Construct

We have said that our tacit-knowledge tests will measure possession of underlying knowledge gained from experience--knowledge identified with Paths A and A' in the explanatory model of tacit knowledge. What we have not yet addressed is the degree to which tacit-knowledge items in the TKS (or the questions on our tacit-knowledge tests) are thought to actually exemplify the content of the underlying knowledge of subjects. This issue is closely related to another issue--the degree to which responses to tacit-knowledge items[30] predict possession of related (but untested) knowledge and performance on domain-relevant tasks.

Do tacit-knowledge items exemplify underlying knowledge? The question here is the degree to which the content of tacit-knowledge items actually reflects the content of the underlying knowledge used by subjects to evaluate or rate the items. We can identify two positions on this issue. The first position is that tacit-knowledge items are exactly the same as the underlying knowledge used to rate or judge them. That is, when a subject endorses a knowledge item we may attribute to her knowledge that is equivalent or isomorphic to the knowledge expressed by the item. From this point of view, tacit-knowledge items are exemplars of the underlying knowledge being measured. A second position is that tacit-knowledge items tell us nothing about the content of underlying knowledge but simply serve to elicit different responses from subjects who differ in their practical, experience-based knowledge. From this point of view, tacit-knowledge items are not exemplars of underlying knowledge--they simply indicate its presence or absence.

Do tacit-knowledge items predict possession of related knowledge and/or performance on related tasks? The question here is the degree to which a subject's endorsement or rating of particular tacit-knowledge items is thought to predict relevant behavior within the domain. Again, we can identify two positions on the issue. The first position is that tacit-knowledge items have predictive value and that this value may be traced to individual differences in the capacity to acquire and use tacit knowledge and the broad effects of those individual differences across tests and tasks. A second position on the prediction issue is, of course, that tacit-knowledge items measure but do not predict--they simply reflect the presence or absence of knowledge sufficient to rate the item (or answer the question) in an experienced or expert-like manner

The two issues outlined above are closely related to a traditional distinction between achievement testing and intelligence testing. In achievement testing, items are presumed to exemplify the measurement construct (e.g., knowledge of world history) but are not commonly viewed as predictors. For example, when a subject correctly answers a factual, multiple-choice question about world history, we assume

---

30 For ease of exposition, the term "item" will be used to refer to both the items of tacit knowledge rated by subjects in the current study and the tacit-knowledge test questions (i.e., scenario descriptions and associated response options) to be constructed from those items.

that she possessed prior knowledge of either the fact in question or related facts that enabled her to rule out incorrect alternatives. We do not commonly view the history question as predictive of performance on other tests or tasks. In intelligence testing, by contrast, items are presumed to predict performance but not are not commonly viewed as exemplars of the measurement construct. For example, when a subject correctly solves a figural analogy problem, we do not assume that he possessed prior knowledge of the analogical relationship in question. However, we do view such analogy problems as predictive of performance on other, "g-loaded" tests and tasks.

Having drawn a distinction between achievement and intelligence testing, it is necessary to point out that neither form of test exists in a pure form. All achievement tests measure underlying abilities--if only the abilities necessary to acquire the tested content--and so tend to have predictive value. Likewise, all intelligence tests measure acculturated knowledge--if only the knowledge necessary to make sense of items and testing conventions--and so tell us something about the knowledge content of individuals rated high and low in general intelligence. Tacit knowledge tests are unique in that they disrespect the traditional (and largely untenable) distinction between achievement and ability testing--they seek both to exemplify content and to predict performance.

Tacit-knowledge tests are knowledge-based tests built on a theory of human intelligence. They are intended to measure both practical, experience-based knowledge and the underlying dispositions or abilities that support the acquisition and use of that knowledge. Thus, scores on tacit-knowledge tests are expected to predict performance on tests or tasks that draw on either tacit knowledge or the mental abilities that supported its creation and use. These abilities are hypothesized to differ from those implicated in the "general factor" in human intelligence commonly referred to as 'g' and measured, in norm-referenced fashion, as IQ. Research by Sternberg and colleagues has produced support for this hypothesis (see "Empirical Research on Tacit Knowledge" above).

We take the position that tacit-knowledge items are both indicators and exemplars of underlying, tacit knowledge and can, at least potentially, shed light upon (1) the content of that knowledge, and (2) the events or experiences through which it was acquired. Few would contest that tacit-knowledge items reflect the knowledge content of subjects from whom they were obtained (e.g., in the course of a "story-telling" exercise focusing on personal experiences). The items came from these subjects' memories and so must reflect the content of those memories. What remains undetermined is the degree to which tacit-knowledge items reflect the content of subjects who did not produce but, rather, endorsed or rated the items. This is not to say that tacit-knowledge items do not serve as exemplars of these subjects' knowledge, but simply that the degree of correspondence is, at present, an empirical question of considerable difficulty.

Of course this is nothing new in the realm of mental testing. A student may answer a multiple-choice question about Vasco de Gama without having any knowledge of that historical figure. By bringing other knowledge to bear, along

with analytical skills, the student can rule out all response options but one and in so doing answer the question correctly. In attributing knowledge of Vasco de Gama to this student we would be incorrect. But in the larger enterprise--assessing the student's knowledge of world history by summing scores on a multi-item test--we would be led to the same (presumably correct) inference. Thus, for the applied purposes of the current project, we choose to treat tacit-knowledge items as exemplars, at an undetermined level of fidelity, of the underlying knowledge of subjects who endorse them. If this assumption introduces error into our analyses we trust that such error will be (1) commensurate with that which is present in a wide range of knowledge-based tests, and (2) more important at the level of individual items than at the level of tacit-knowledge tests or test batteries.

Supporting Instrument Development

Because tacit-knowledge tests are, in a sense, hybrids of achievement tests and ability tests, they differ somewhat from either of these types of tests in the way in which they are constructed and validated. In achievement testing, content validation takes precedence over construct or criterion validation--the content is the construct and mastery of content is the criterion. In tacit-knowledge testing, there are no objectively correct answers and so the measurement of concurrence with an expert response profile is intrinsic to test scoring. In intelligence testing, measurement of concurrent validity has traditionally predominated over the evaluation of content. An item or class of items that loads heavily on the general factor for human intelligence is deemed to measure the underlying construct.[31] In tacit-knowledge testing, however, a theory about human-knowledge acquisition specifies what counts and does not count as tacit knowledge. For this reason, a strictly correlational approach to item selection is undesirable.

By examining the relationship between the goodness ratings applied to tacit knowledge items and the experience/effectiveness of the leaders giving the item ratings, we created indices of promisingness that will be useful in selecting items for incorporation into tacit-knowledge tests. However, these relationships to external criteria are not the only, or even the most important, indices of promisingness in the test-construction effort we envision. The reader will notice, in this regard, that we did not seek to identify items that correlated significantly with our criteria but trivially with each other, as is sometimes the practice in test development. As Nunnally (1970) and others have argued, such a "criterion-based" approach to test development is problematic and often produces measurement instruments of inferior quality. Specifically, such an approach may be expected to yield tests that suffer

---

[31] There are, of course, notable exceptions to this characterization in the recent history of intelligence research (see Sternberg, 1990).

from low internal-consistency reliability, poor factor structure, and fragility with respect to criteria other than those on which the selection of items was based.[32] Thus, rather than "chasing a criterion" we intend to select items that best measure the tacit-knowledge construct, basing our eventual decisions on a combination of evidence (e.g., empirically demonstrated relationships) and rhetoric (e.g., arguments for the face validity of items with respect to the focal construct).

A Unified Approach to the Validity of Tacit-Knowledge Tests

In order to show how the results of the current study come together to support the goal of construct measurement, and in order to set the stage for the development effort to come, we now turn to an enumeration of the various aspects of validity in psychological measurement. We employ a unified validity framework, set forth by Messick (1995), that treats the traditionally separate forms of validity (i.e., content, construct, and criterion) as aspects of a more comprehensive construct validity. According to this framework, the essential goal of test validation is to support, through a combination of theoretical rationale and empirical evidence, the interpretation of tests scores and the uses of scores under that interpretation.

The content aspect.

The content aspect of validity refers to evidence that test content is relevant to and representative of the focal construct.[33] In the context of tacit-knowledge test development, the goal of construct relevance calls for tacit-knowledge test questions that are sensitive to knowledge of the type specified by the focal construct and insensitive to knowledge that falls outside the focal construct. A first step towards this goal was taken during the identification phase of the project when, in interviews with Army officers, we oriented subjects toward personal experiences and away from leadership doctrine or theory. A second step was taken, in the current study, when we obtained goodness ratings on tacit-knowledge items. These ratings (i.e., question means and variances) may serve as a source of evidence regarding the relevance of tacit-knowledge items to the underlying construct. For example, questions with low mean goodness ratings and low variances are items with a poor claim to having been learned through personal experience, given the size and breadth of the sample of officers rating the items. In the future, we may wish to further pursue the goal of construct relevance by routing candidate test questions to military experts, along with a simply worded explanation of the focal construct, in order to elicit judgments of relevance.

---

[32] A purely criterion-based approach would be problematic in tacit-knowledge test construction because test items are scored with reference to an expert response profile that (depending upon how experts are selected) may reasonably be expected to correlate with criterion measures.

[33] The content aspect includes concerns that traditionally fall under the heading of "content validity."

The goal of construct representativeness calls for tacit-knowledge items that are typical rather than atypical of knowledge items specified by the focal construct. An important step toward this goal was taken during the identification phase when an expert panel eliminated from consideration knowledge items judged to be technical in nature in favor of knowledge items that addressed leadership as an influence processes. A second step was taken, in the current study, by asking subjects to rate knowledge items on the "often" scale. Thus, for example, questions with low means and variances on the often scale may be those with the poorest claim to representativeness. Finally, the scaling analysis of sort data produced a tentative model of the tacit-knowledge space (albeit a model limited by properties of the knowledge sample) and each items generalized distance from the origin of this space may serve as a rough index of typicality and representativeness. In the future, we may wish further to pursue the goal of construct representativeness by routing candidate test questions to military experts, along with a simply worded explanation of the focal construct, in order to elicit judgments of typicality.

The substantive aspect.

The substantive aspect of validity refers to the theoretical rationale embodied in an explanatory model of task (test) performance. A major step toward the goal of substantive validity was taken, in the current report, when an explanatory model of tacit knowledge was proposed and elaborated. The explanatory model distinguishes tacit knowledge from nontacit job knowledge by associating the two forms of knowledge with different knowledge-acquisition pathways. The explanatory model makes straightforward predictions about the performance benefits of tacit knowledge in a range of task situations. Importantly, the explanatory model of tacit knowledge predicts that the possession of tacit knowledge will confer an advantage (relative to that conferred by nontacit job knowledge) in responding to contextualized problems of realistic complexity. Thus, the explanatory model of tacit knowledge constitutes a high-level model of tacit-knowledge test performance and, as such, directly serves the goal of substantive validity. In future work, we may wish to collect empirical evidence regarding the explanatory model of tacit knowledge. Specifically, we may wish systematically to collect self-report data from subjects taking tacit-knowledge tests. In so doing, we may be able to obtain evidence bearing on the degree to which subjects draw on personally experienced, rather than received, knowledge of military leadership.

The structural aspect.

The structural aspect of validity refers to the level of fit between the internal structure of the test and the internal structure of the construct domain. A first step toward the goal of structural validity was taken, during the identification stage of the project, when we interviewed and elicited knowledge from Army officers in all three of the major branch categories within the Army (i.e., combat arms, combat support, combat service support). The goal of structural validity was further served, in the current study, when we collected TKS data in a wide variety of TRADOC facilities and FORSCOM units--literally, from rifle companies to maintenance battalions. In short,

by using broad samples of Army officers, we sought to avoid basing our analyses and test development on a restricted subset of leadership tacit knowledge. Of course, the structural aspect of validity is addressed, most directly, by the scaling analyses performed on sort data and by the knowledge dimensions derived in that analysis. Having sought to cast a wide net in our sampling of tacit-knowledge items, we have examined the internal structure of that sample--as reflected in the sort data of our USMA subjects. In so doing, we have improved our prospects for developing tacit-knowledge tests that mirror the structure of the construct domain (i.e., the population of practical, action-oriented knowledge that Army leaders acquire from personal experience).

The generalizability aspect.

The generalizability aspect of validity refers to the extent to which score properties and interpretations generalize across groups, settings, and tasks.[34] In the context of tacit-knowledge test development, the goal of generalizability calls for tacit-knowledge tests for which score interpretations generalize across (1) roles within the organization, (2) repeated administrations, and (3) alternate forms of the test. Although these are primarily concerns to be addressed in test development, efforts relevant to the content, substantive, and structural aspects of validity are also relevant to the generalizability aspect. In general, by seeking to specify and measure the construct, rather than merely pursuing correlation with an external criterion, we presumably increase the generalizability of score interpretations for our tacit-knowledge tests.

The external aspect.

The external aspect of validity refers to convergent and discriminant evidence from multi-trait/multi-method comparisons as these speak to the overarching issue of construct validity.[35] In the context of tacit-knowledge test development, possible convergent evidence would include correlation between tacit-knowledge test scores and exogenous variables such as rated leader effectiveness, degree and rate of career advancement, and performance on construct-relevant tasks. Again, our efforts to specify and measure the construct provide the most important support for this goal. In addition, results of discriminant and correlational analyses from the current study will provide a basis for evaluating items in terms of their potential contribution to the external validity of tacit-knowledge tests.

In the context of tacit-knowledge test development, possible discriminant evidence would be that which discounts the effects of general intelligence, reading comprehension, and general job knowledge on tacit-knowledge test scores and on the convergence of these scores with external indices of performance. In earlier work,

---

[34] The generalizability aspect includes concerns that traditionally fall under the heading of "reliability."

[35] The external aspect includes concerns that traditionally fall under the heading of "criterion validity."

on tacit knowledge for managers, discriminant evidence was obtained by entering tacit-knowledge test scores into an hierarchical regression analysis along with scores on other psychological tests, including a measure of general ability. In pursuing similar analyses for our military sample, the explanatory model of tacit knowledge may provide a basis for more clearly contrasting tacit and nontacit job knowledge, in order to further establish the incremental validity of tacit-knowledge test scores.

The consequential aspect.

The consequential aspect of validity refers to the value implications of the intended use of score interpretation as a basis for action. Tacit-knowledge tests are not intended, or commonly used, for employment selection. Thus, many of the value implications traditionally included under the consequential aspect are not relevant. However, because the development of tacit-knowledge tests is intended for use in leader development efforts within the Army, it will be important to consider the degree to which construct-relevant knowledge that contradicts Army doctrine or culture is appropriate for incorporation into tacit-knowledge tests.

Conclusion

The goal of the current study was to support the development of tacit-knowledge tests for which the interpretation of scores will be valid. Our description of a high-level explanatory model of tacit knowledge constitutes, we believe, a major step in the direction of this goal. By elaborating the tacit-knowledge construct at the cognitive level, we set the stage for a more detailed consideration of item content during the selection process and, in so doing, increase the substantive validity of our tests. The analysis of item and leader rating data constitutes a second step towards measuring the construct. By identifying those items with the strongest association with performance criteria, we increase the probability that we will select items and construct test questions that embody the construct--given that the construct model makes clear predictions about the performance benefit of tacit knowledge. The analysis of latent structure in sorting data constitutes a third step towards our goal. By examining the structure of the tacit-knowledge space (based on our sample) we will be able to make more informed decisions about the distribution of item content in our tacit-knowledge tests and, in so doing, increase the structural validity and generalizability of score interpretations. Finally, by relating our results to a unified validity framework, we have set the stage for the content selection, question construction, test construction, and test validation that will be the new focus of work in our ongoing effort to understand and facilitate the process of leadership learning in the military.

References

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Broadbent, D. E., & Aston, B. (1978). Human control of a simulated economic system. Ergonomics, 21, 1035-1043.

Broadbent, D. E., Fitzgerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. British Journal of Psychology, 77, 33-50.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). The nature of expertise. Hillsdale, NJ: Erlbaum.

Groen, G. J., & Pate. V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), The nature of expertise. Hillsdale, NJ: Erlbaum.

Hartigan, J. A. (1975). Clustering algorithms. New York: Wiley.

Hays, W. L. (1988). Statistics. New York: Holt, Rinehart, and Winston.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). Induction: Processes of inference, learning, and discovery. Cambridge, MA: MIT Press.

Horvath, J. A., Forsythe, G. B., Sweeney, P. J., McNally, J. A., Wattendorf, J., Williams, W. M., & Sternberg, R. J. (1994b). Tacit knowledge in military leadership: Evidence from officer interview (Technical Report 1018). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A289 840)

Horvath, J. A., Williams, W. M., Forsythe, G. B., Sweeney, P. J., Sternberg, R. J., McNally, J. A., & Wattendorf, J. (1994a). Tacit knowledge in military leadership: A review of the literature (Technical Report 1017). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A291 140)

Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 21-38.

James, W. (1890). The principles of psychology (Vol. 1). New York: Dover.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons; responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-750.

Neisser, U. (1976). Cognition and reality. San Francisco, CA: Freeman.

Nunally, J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill.

Polanyi, M. (1966). The tacit dimension. Garden City, NY: Doubleday.

Reber, A. S. (1967). Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 6, 317-327.

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology, 81, 115-119.

Reber, A. S., & Millward, R. B. (1968). Event observation in probability learning. Journal of Experimental Psychology, 77, 317-327.

Schacter, D. L. (1987). Implicit memory: History and current status. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 501-518.

Schmidt, F. L., & Hunter, J. E. (1993). Tacit knowledge, practical intelligence, general ability, and job knowledge. Current Directions in Psychological Science, 2(1), 8-9.

Schon, D. A. (1983). The reflective practitioner: How professionals think in action. New York: Basic Books.

Sternberg, R. J. (1990). Metaphors of mind: Conceptions of the nature of intelligence. Cambridge: Cambridge University Press.

Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and school. In H. W. Resse and J. M. Puckett (Eds.), Mechanisms of everyday cognition. Hillsdale, NJ: Erlbaum.

Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. American Psychologist.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press.

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. Journal of Personality and Social Psychology, 52, 1236-1247.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. Journal of Personality and Social Psychology, 48, 436-458.

Williams, W. M., & Sternberg, R. J. (in press). Success acts for managers. New York: Harcourt Brace.

Appendix A

Implicit Memory: A Typical Demonstration

During the training phase of a memory experiment, subjects are presented with a large number of words, one at a time on the computer screen. Subjects are asked to watch for the occurrence of a particular letter sequence. In a later testing session, subjects are shown a new list of words and asked to identify those which they saw during training. When the length of the training list is long, subjects typically fail to remember having seen many of the words that were present on the training list. Thus, in a test of recognition memory, subjects show forgetting of many of the individual words on the list. When memory for individual words is tested in other ways, however, the presence of latent or implicit memories for individual words can be demonstrated.

In a typical procedure, subjects are presented with lists of text strings and are asked to decide, as quickly as possible, whether or not the presented text string is a legitimate word in English or a nonsense word. Using carefully constructed stimulus sets, experimenters have been able to show that test words that are semantically related to words seen during the training phase are responded to more quickly than are test words unrelated to words seen during the training phase. The accepted explanation for this effect is that implicit memory for the words seen during training facilitates or "primes" responses to the semantically-related test words. The ability of "forgotten" words to prime responses to semantically-related words is taken as evidence for the existence of implicit memory for these words--memories that subjects did not know they possessed.

Appendix B

The Relationship between Tacit Knowledge and Proceduralized Skill

In this appendix we remark on some of the salient similarities and differences between tacit knowledge (as conceptualized in our explanatory model) and proceduralized skill (as conceptualized in the work of Anderson, 1983).

First, although tacit knowledge and proceduralized skill are both forms of knowledge that may be based on personal experience, they represent different types of information about the world. The domain of tacit knowledge is essentially the covariance structure of the environment--what goes with what in the world. The domain of proceduralized skill is patterned, complex, goal directed behavior--how to do things in the world. Second, although both tacit knowledge and proceduralized skill are tacit or opaque to conscious awareness, their tacitness takes different forms. Tacit knowledge is tacit in its acquisition (people do not realize what, or that, they are learning) but it may become conscious through reflection. Proceduralized skill, by contrast, is nontacit in its acquisition (when action sequences are being deliberately practiced) but it may become tacit through automatization.

Finally, although both tacit knowledge and proceduralized skill can be modeled with if-then rules, the nature of the consequent or "then" portions of their rules are different. Tacit knowledge can be modeled by rules in which consequents represent classification decisions (IF <it quacks> AND <it waddles> THEN <it's a duck>) or describe a transition in the world (IF <it falls off the counter> AND <the kitchen floor is tile> THEN <it will shatter>). Proceduralized skill can be modeled by rules in which the consequents represent actions to be taken by the thinker (IF <you like her> AND <she smiles at you> THEN <go up and talk to her>).

Contrasting the three types of if-then rules (i.e., categorical, causal, and prescriptive) suggests that declarative knowledge and proceduralized skill interact and complement each other (see Anderson, 1983, p.215). In acquiring a procedural skill (e.g., how to conduct a performance review) one requires information about how the world works during the period of acquisition--when action sequences are being assembled and tried out. That is, one needs information about categorical relationships in the world (e.g., "an OER is one type of performance evaluation") and about causal relationships in the world (e.g., "if you emphasize status differentials you may cut off honest communication"). Both types of information support the development of new skills. Thus, declarative knowledge of both a tacit nature (based on personal experience) and a nontacit nature (based on received wisdom) can support the development of proceduralized skill.

The influence of tacit knowledge on proceduralized skill can be seen as an important mechanism through which tacit knowledge has an impact on behavior/performance. A key to the human capacity for sophisticated, goal-directed behavior is the ability to automatize well-learned, behavior sequences and, in so doing, to free conscious awareness for processing additional information. Thus, one of the ways in which tacit knowledge manifests itself in behavior is by supporting

the development of proceduralized skills. For example, declarative knowledge about the way in which status differentials affect communication can form the basis for a sequence of actions that minimize the effects of status differentials and, thus, promote open communication.

When we see tacit knowledge in the world, it is often packaged in the form of proceduralized skill. Does this mean that tacit knowledge and proceduralized skill are the same thing? According to our current thinking, the answer to this question is no. We take the position that tacit knowledge and proceduralized skill are complementary but, in theory, distinguishable. The fact that tacit knowledge often manifests itself in the form of well-practiced behavior sequences does not imply that tacit knowledge is reducible to proceduralized skill--any more than the fact that love is often expressed through valentines means that love is reducible to valentines. A valentine is one of an arbitrary number of ways in which love can be expressed through behavior. Likewise, a proceduralized skill is one of an arbitrary number of ways in which declarative knowledge about the world can be expressed through behavior.

Thus, because proceduralized skill may incorporate both tacit knowledge (based on personal experience) and nontacit knowledge (based on received wisdom), the tacit knowledge <u>construct</u> cannot be based on the condition-action structure that has been used to model proceduralized skill. There is no reason, however, why actual tacit knowledge cannot be <u>expressed</u> in the form of procedures, as we have done in some of our prior work.

Appendix C

Instructions to USMA Judges

On behalf of my colleagues at Yale and at West Point let me thank you for making the time to help us with this aspect of our research. The dimension-interpretation task that you are undertaking is an intellectually demanding one. It will require that you draw on your practical knowledge and personal experience of military leadership, as well as on your more general skills in analysis, problem solving, and hypothesis testing. Please know that we have not subjected you to these challenges (or asked for your valuable time) without careful consideration. It is our hope that your effort on this task will contribute materially to an improved understanding of Army leadership and the process of leader development through job assignment.

Major Bullis will instruct you in the basic nature of the task and will answer any questions that you may have. Let me close by offering a few guidelines that we hope you will follow in interpreting the various dimensions of leader knowledge with which you will be presented. Note that these guidelines may not make much sense until Major Bullis has described the task to you and, perhaps, until you have had a look at the materials.

GUIDELINES FOR JUDGES:

First, please make every effort to divest yourself of prior theories or classification schemes in interpreting these dimensions. As much as possible, let the dimension labels emerge from the knowledge items themselves, rather than from theories or taxonomies of leadership behavior or roles that you may have been exposed to. That is to say, take a "data driven" rather than "theory driven" approach to the task.

Second, do not expect the dimension labels to come easily or quickly. The knowledge items that make up each dimension are highly complex and should require a good deal of reflection, hypothesis testing and, eventually, discussion. Do not be surprised if you come up with several alternative labels for a given dimension and have difficulty deciding among these.

Finally, you may find it useful to use a method of contrast wherein items that have clustered together on one dimension are examined for their (a) position and (b) dispersion on other dimensions. This contrast method, which is easier to use than to describe, can help you to isolate the commonalties and differences that underlie the various dimensions.

Again, we appreciate your help in this matter and will do our best to make sure that the interpretations you generate are put to good use in our on-going effort to understand and promote the development of Army leaders.